The ATLAS SCT alignment system and a comparative study of misalignment at CDF and ATLAS

Stephen M. Gibson

St. John's College

University of Oxford

Thesis submitted in partial fulfilment

of the requirements for the degree of

Doctor of Philosophy

University of Oxford, Hilary Term 2004.

The ATLAS SCT alignment system and a comparative study of misalignment at CDF and ATLAS

Stephen M. Gibson
St. John's College, University of Oxford

Thesis submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy.

University of Oxford, Hilary Term 2004.

Abstract

A promising new frontier in particle physics will soon be accessible at the Large Hadron Collider. The ATLAS experiment will exploit this opportunity to explore the nature of electroweak symmetry breaking and search for physics beyond the Standard Model. The accurate alignment of the ATLAS tracking detector will be crucial for appropriate sensitivity to discovery and precision physics.

The challenge of aligning the colossal tracking detector of ATLAS has stimulated the development of a novel alignment system, based on a geodetic grid of over eight hundred length measurements between nodes attached to the detector support structure. This thesis contributes to these developments in three main ways:

- Important improvements to the reliability, precision and scope of the system's essential interferometric length measurement technique (FSI) have been demonstrated. A 170 mm stabilized interferometer has been measured to 7 nm. The precision achieved with prototype interferometers of the same design as for ATLAS, is a factor of 5 better than required.

- The first prototype FSI grids of multiple, simultaneous length measurements have been constructed and rigorously tested. A method of precisely reconstructing the shape of these grids has been developed. This method has been verified with real measurements by exploiting the redundancy in the grid.

- Detailed simulations of the ATLAS SCT alignment system predict it should be capable of reconstructing the SCT shape to within the required precision. A grid design for the SCT End-cap has been developed.

Additionally, misalignments in CDF II were evaluated and mapped onto ATLAS to provide an indicative estimate of ATLAS commissioning misalignments. These studies predict misalignment induced momentum degradation will be significant for $p_T \gtrsim 30$ GeV/c^2 and the b-tagging performance will be degraded by at least 10%. The SCT alignment system will play a vital role in ATLAS.

To my wife, Jennifer, and my parents.
For all your loving support and encouragement.

My soul will boast in the LORD;
Let the afflicted hear and rejoice.
I sought the LORD, and he answered me;
he delivered me from all my fears.

Psalm 34: 2,4.

This thesis was typeset with LaTeX 2_ε.

Published at the University of Oxford, Oxford, United Kingdom.

Acknowledgements

I am delighted to be able to thank all those who have made these last $3\frac{1}{2}$ years (to the day!) not only instructive but pleasurable.

Heart-felt thanks are deserved for my supervisor, Tony Weidberg, not only for his generosity with tea and apple tartes, but also for his "strange idea" behind the physics analysis in this thesis and for all his friendly wise words and encouragements. The other ATLAS academics deserve thanks: Richard Nickerson has paced considerably faster this year and credit is due for his bearing of the political onslaught while the first part of the FSI system was assembled. I am particularly grateful to David Howell, for many interesting and fruitful discussions, especially early on, about hardware and electronics, and for enlightening me about four-pole Butterworth filters. And I am also grateful for his help when the laser, erm, *happened* to blow up. Thanks also to Todd Huffman for enthusiastically telling me about retro-punching and the WHIP-eM system (please don't get ideas for the viva!) and also to Armin Reichold for enjoyable discussions on varied topics that reveal his contagious enthusiasm for physics.

Now for my fellow students in the ATLAS office. I feel honoured to have been among so many fine upstanding gorillas. Thanks to Ankush Mitra, for his infallible help with electronics and much good humour. Thanks to, Antonia Lavorato-Cioffi, who not only has the good sense to be from Italy, but also to have been married there and invite us all to join her - thank you for much lively conversation, Italian cuisine and warm friendship. It has been a great pleasure getting to know Danny Hindson, to ponder physics we really ought to know together and to share in his frustration with the CERN batch system. Thanks also to Alex and Jonathan, Mohamed and Nick, for their growing concern for me over the past few months: hope to see you all at the pub.

Various busy people have spared the time to chat with me about physics and alignment. Thanks to Stephen Haywood, Alan Poppleton, Dario Barberis and especially to Jose Enrique Garcia for providing his b-tagging code. Special thanks are deserved for Thilo Pauly who provided the CDF II data for this thesis and who helped me to decipher it. Thanks also to Farrukh Azfar and Jonas Rademacker for helpful comments. I am also grateful to Laurent Brunel, for kind permission to use the SIMULGEO software crucial to this thesis and for teaching me its tricks. I would also like acknowledge Greg and Helen Heath for inspiring my interest in particle physics.

Thanks to those in the Oxford physics drawing office who have created designs for the FSI work: Wing-commander-Lau, Brian O., Nick B., Jim L. and Stephanie Y. The workshop team deserve a special mention for endlessly punching retro-reflectors and for making all the FSI bits and pieces and the pro-

totype grid components: Tony Handford, Mike M., Mike T., Charlie, Richard and Murray among others. I am particularly thankful for Paul Pattinson's friendly help with essential optics and cables, and also for his helpful replacement, Richard Makin. I am also grateful for all the support staff in physics, especially to Pete Grönbech for keeping pplxgen running and to John Harris for help with urgent printing of this thesis.

I gratefully acknowledge PPARC for the financial support to start this work and for providing the deadline by which to finish it.

I reserve my most grateful thanks for Paul Coe, whose help over past few years has been immeasurable. While spending many entertaining months together in the FSI lab, writing LabView, injecting glue or puzzling over the finer details of FSI, we got to know each other rather well. Our naturally opposite approach to problem solving has often lead to animated discussions and the best results. He has taught me how best to investigate a problem and do research, honed my linguistics, and stretched my understanding of FSI to breaking point. Above all he has been my friend. I am very grateful for his meticulous reading of the majority of this thesis and for his encouragements throughout my studies.

When the physics is done, I am grateful to all those who have helped make these past few years enjoyable. Despite the apPauling puns, Paul has often lead the charge, with the likes of pancake and Feuerzangenbole evenings. I must mention those two friendly, Forward people: Pawel Bruckman and Sandra Muijs (soon to be Mice). I will not list all my friends at Oxford and elsewhere but you are all invited to the potato party.

Finally to those who have greatly encouraged me:

I praise God who has sustained me throughout and for friends at WRBC and elsewhere, for their prayerful support. Thank you, Mark, for encouraging me to finish with the lure of a tweed jacket. I am very grateful too for my parents continual enthusiasm and encouragement. Thank you, Jennifer, for lovingly supporting me in so many ways. Although in these past few months I was not often with you, I could not have done it without you.

Thank you everyone.

Preface

The Large Hadron Collider presents an exciting opportunity to probe the nature of electroweak symmetry breaking and to search for New Physics. The challenging LHC environment places extraordinary demands on the detector technologies and on the alignment of suitable tracking detectors. A novel alignment system for the ATLAS SemiConductor Tracker has been developed that will rapidly enhance the sensitivity of ATLAS to discovery and precision physics. Movements of detector modules will be reconstructed in quasi real time from over eight hundred length measurements between nodes attached to the tracking detector. This thesis contributes to these developments.

Chapter 1 A brief review of current progress in particle physics, developments for ATLAS and the motivation for the SCT alignment system.

Chapter 2 An introduction to the precise length measurement technique developed for the ATLAS SCT alignment system: Frequency Scanning Interferometry. Crucial advances in the technique that improve the reliability, precision and scope of the measurements are presented. The FSI measurements are assessed for tolerance to spurious effects that may degrade the interferometer signals in ATLAS.

Chapter 3 For the first time, geodetic grids of multiple FSI measurements have been constructed and measured. This chapter describes the experimental set up of these prototype grids and important issues regarding optimization of the interferometer signals.

Chapter 4 Shape deformations of the ATLAS SCT will be precisely reconstructed from FSI grid measurements. In this chapter, the grid shape reconstruction technique is developed using measurements of the prototype grids of Chapter 3. The first demonstration of precise grid shape reconstruction with FSI is presented.

Chapter 5 This chapter presents the simulated performance of the FSI grids that will monitor the ATLAS SCT. Development of the SCT End-cap FSI grid is outlined.

Chapter 6 Analysis of misalignments at CDF II has provided an indication of the commissioning misalignments that can be expected at ATLAS. These studies are presented and the implications for ATLAS physics are investigated. The studies emphasize the vital role of the ATLAS SCT alignment system.

Chapter 7 The main conclusions are summarized.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Standard Model and Beyond

The Standard Model [SM] of particle physics has emerged from the twentieth century theoretical and experimental endeavours to understand the fundamental constituents of matter and their interactions. The Standard Model is a spontaneously broken, renormalizable, quantum gauge field theory, that has successfully described the phenomenology of almost all high energy particle physics, since the theory's formulation in the 1970's. The predictions of the intermediate vector bosons, $W^{\pm}$, Z^0, and the top quark were a considerable success of the theory. Its validity in the electroweak sector has been verified by LEP during the past decade with unprecedented accuracy.

Despite the success of the Standard Model, it cannot be the final theory. The inadequacies of the Standard Model and the future agenda for particle physics may be considered under five categories (after [Quig02]):

Elementarity The constituent quarks and leptons seem elementary at the current limits of resolution ($\sim 10^{-18}$ m), but is there an underlying structure that might explain their properties?

Symmetry Nöther's theorem has proved a powerful tool to elucidate particle interactions from symmetry properties. Which gauge symmetries exist and why? How is the electroweak symmetry hidden? The search for a more fundamental theory is largely a search for higher symmetry principles.

Unity The disparity between the strengths of the interactions is large, at the energies accessible by current accelerators. Might the gauge couplings unify near to the Planck scale? Quarks and leptons appear to be structureless, spin 1/2 particles: (how) are they related? Why are there at

least three generations? Why is charge quantized? How can gravity be incorporated? Various Grand Unified Theories seek to address these questions, some with considerable success, but which model matches reality? An experimental lead is needed.

Identity Recent observations of neutrino oscillations [SuperK, SNO] necessitate the first significant change to the Standard Model. What causes neutrinos to apparently change identity? What determines the phase in the quark-mixing matrix that describes CP violation?

Topology What is the fabric of space-time? Are there large extra dimensions?

An appealing solution to some of these problems is supersymmetry (SUSY), discussed below. First, the current experimental status of the last, infamously missing piece of the Standard Model is reviewed.

The Higgs boson

Weak isospin is not an exact symmetry due to the difference in mass within the fermion doublets and between the masses of vector bosons. Attempting to add mass terms to the electroweak Lagrangian destroys the gauge invariance and results in a non-renormalizable theory. The Higgs mechanism [SM] solves this problem by introducing a complex doublet of scalar fields which have a scalar potential with a degenerate ground state. In this mechanism, the $SU(2)_L \otimes U(1)_Y$ symmetry is *spontaneously broken* by the choice of ground state about which to perform perturbative calculations. The mechanism gives mass to the $W^{\pm}, Z^0$ bosons, leaving the photon as a massless neutral boson of an unbroken $U(1)_{em}$ symmetry. The remaining degree of freedom is attributed to the Higgs mass: a free parameter of the Standard Model that is yet to be directly measured.

The existence of a Higgs boson or similar object(s), is also indicated by the precision measurements of the $e^+e^- \rightarrow W^+W^-$ cross section at LEP, as in Figure 1.1. The calculated cross section fits the data only when the tree level diagrams with trilinear gauge couplings, WWγ and WWZ^0, supplement the neutrino exchange diagram. These additional diagrams lead to unitarity violating divergences at high energies, which must be cancelled either by the introduction of the Higgs channel or some other underlying mechanism. Similar unitarity constraints require the Higgs mass be less than $\left(\frac{8\pi\sqrt{2}}{3G_F}\right)^{\frac{1}{2}} \approx 1$ TeV [Quig02].

A light Higgs boson is favoured by the current M_W-m_t-M_H correlation, as shown in Figure 1.2 and by global fits to electroweak data, as in Figure 1.3, when analysed *within the Standard Model framework*. The 95% CL upper

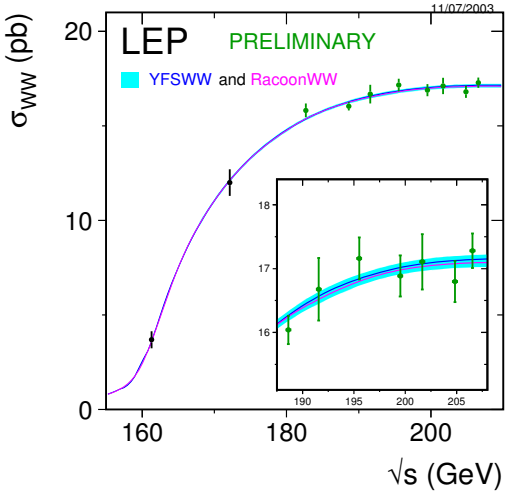

Figure 1.1: *Cross section for $e^+e^- \to W^+W^-$ measured by the four LEP experiments. [LEPEWG]. Trilinear gauge couplings must be included to match the data.*

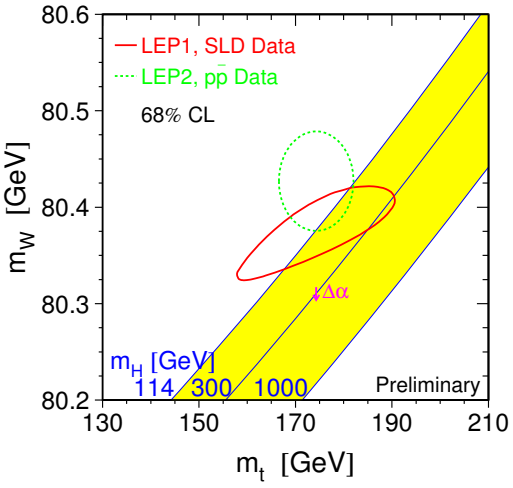

Figure 1.2: *Indirect (LEP 1, SLD) and direct (LEP 2, Tevatron) measurements of M_W and m_t. The mass relationship as a function of the Higgs mass is shown. [LEPEWG]*

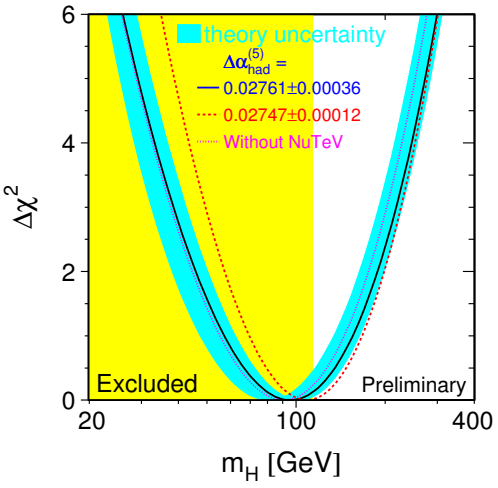

Figure 1.3: *Indirect and direct experimental limits on the Higgs showing the χ^2 of electroweak precision fit and the 95% CL exclusion limit from direct searches at LEP. [LEPEWG]*

limit including theoretical uncertainty is $M_H \lesssim 219$ GeV and direct searches conclude $M_H > 114.4$ GeV [LEPEWG], so either the Higgs should soon be discovered or the Standard Model analysis is misleading.

Supersymmetry

One of the most promising extensions to the Standard Model is supersymmetry (SUSY), which introduces a fermion-boson symmetry and postulates supersymmetric partners for all the known particles. Approximately doubling the particle spectrum does not help reduce the number of free parameters, but SUSY does have several attractive consequences. In the Standard Model, quantum loop corrections to the squared mass of an elementary Higgs boson diverge quadratically with the cut off scale, Λ [Ell98]. SUSY exploits the opposite signs of boson and fermion loops to cancel the divergence and avoids artificial fine tuning. SUSY also suggests convergence of the gauge couplings at high enough energies and provides a mechanism for incorporating gravity.

1.2 The LHC and ATLAS

1.2.1 The Large Hadron Collider

The significance of the TeV scale to searches for the Higgs and New Physics, and the energy limit of the LEP synchrotron, have stimulated the construction of the Large Hadron Collider (LHC) facility at CERN. Counter-rotating beams of protons will collide with unprecedented energy and luminosity providing abundant event rates for rare physics channels. The 14 TeV proton-proton centre-of-mass energy should provide several TeV to the hard physics process. The luminosity should initially be $\mathcal{L}_{low} = 10^{33} cm^{-2} s^{-1}$, and then be increased to the design luminosity, $\mathcal{L}_{high} = 10^{34} cm^{-2} s^{-1}$. Proton bunches will cross every 25 ns to produce about 23 interactions per crossing at design luminosity. First collisions are scheduled for 2007.

1.2.2 The ATLAS Detector

Overview and physics aims

The inherent QCD background at a hadron collider presents a challenging environment from which to glean interesting physics. The LHC particle detectors must have fast response times, fine granularity and be radiation hard. The ATLAS detector [ATL99], shown in Figure 1.4, has been designed to meet these stringent requirements. It is a general purpose detector that enables

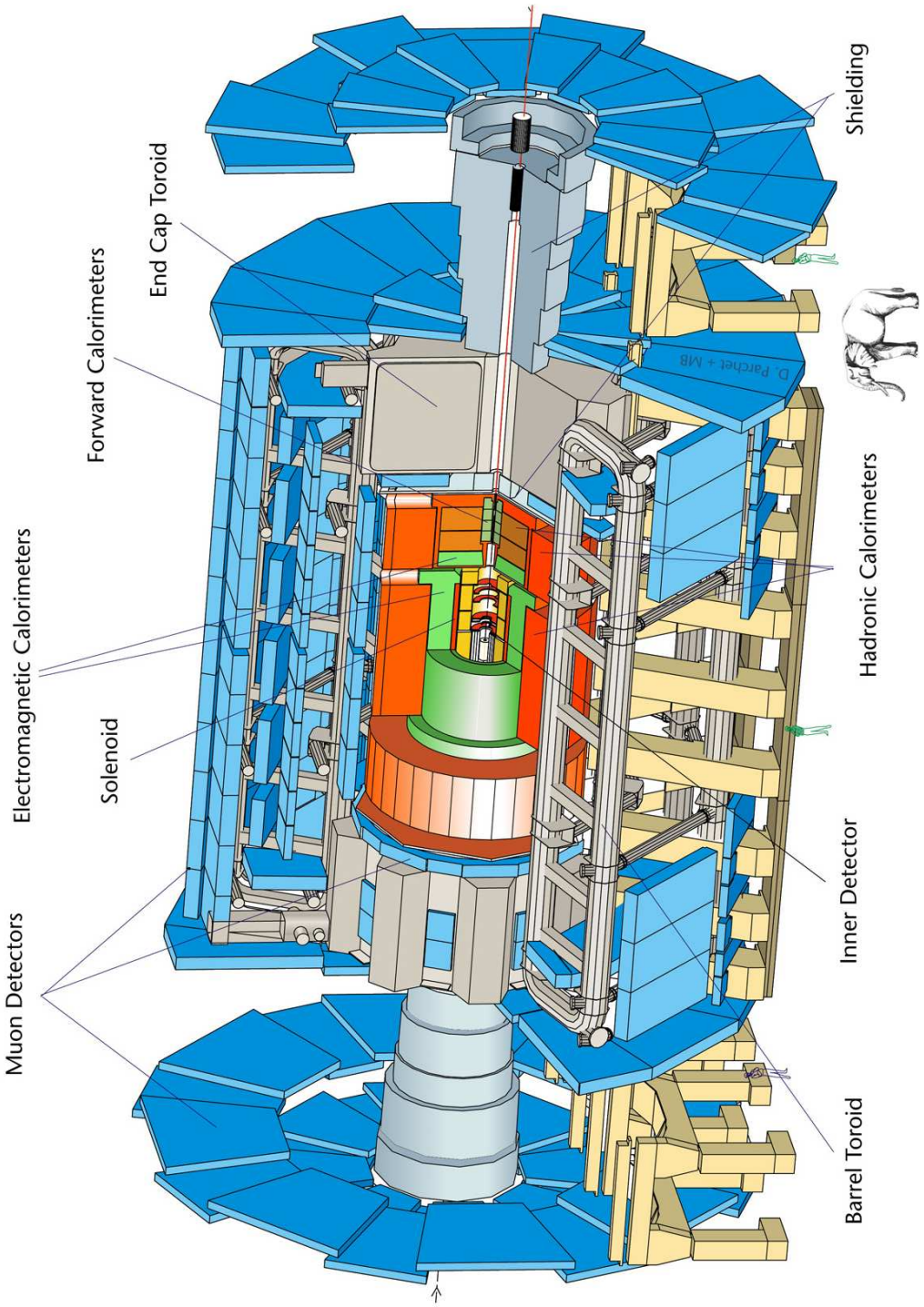

Figure 1.4: *A Toroidal LHC ApparatuS: 44 m long, 22m high and 4000 tons [ATL99]. The elephant is African.*

many known physics channels to be precisely measured and many new physics channels to be searched for:

Higgs boson(s) Elucidating the mechanism of electroweak symmetry breaking is the priority for the LHC. $H \rightarrow ZZ^* \rightarrow 4l$ is the most promising Higgs channel in ATLAS for $130 \lesssim M_H \lesssim 180$ GeV.

SUSY searches Supersymmetry signatures, often characterized by missing E_T and a lepton tag, should be clearly observable in ATLAS.

Top physics The LHC will be a top factory allowing precise measurements of the top mass and interesting top phenomenology studies (e.g. $t \rightarrow H^+b$). Excellent b-tagging is required.

B-physics CP violation in the B-sector should be significant. Although ATLAS lacks the π/K separation of LHCb [LHCb03], the copious b-quark production rate can still be exploited by ATLAS, through b-tagging and soft lepton tags.

The most exciting prospect is that of unexpected physics. Configuring ATLAS to difficult benchmark signatures should enhance sensitivity to any new channels.

The data rate at ATLAS must be reduced for feasible readout and storage. The numerous raw events are filtered by a three level trigger system. The aim is to remove minimum bias events from the interesting physics. The 1 GHz interaction rate is reduced to 100 kHz by the level 1 trigger, with events placed in a pipeline pending the trigger decision. The second level trigger uses full granularity to examine regions of interest identified by level 1. After the level 3 filter and event builder, the storage rate is 100 Hz, amounting to 1 Pbyte per year of running. Computational *grids* are being developed to handle the data analysis [Fos99].

Only the ATLAS Inner Detector, which is relevant to this work, is outlined below with an emphasis on recent developments. Detailed descriptions of all the sub-detectors shown in Figure 1.4, are provided elsewhere [ATL99].

Inner Detector

Charged particles emerging from the interaction point with pseudorapidity[1] $|\eta| < 2.5$ are tracked by the ATLAS inner detector, shown in Figure 1.5, which has three elements:

[1] $\eta \equiv -\ln\tan(\theta/2)$, where $\cos\theta = p_z/p$ defines the production angle θ with respect to the beam axis.

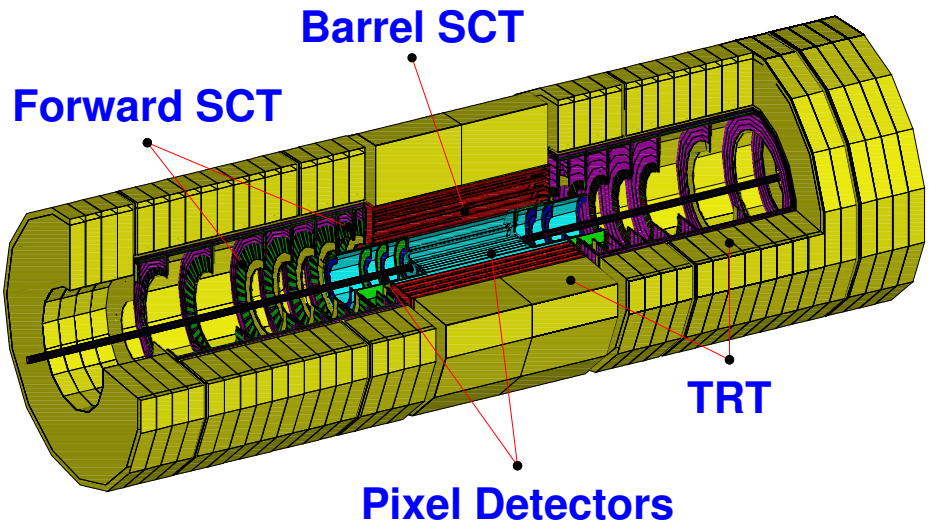

Figure 1.5: *The ATLAS Inner Detector is 7 m long with an outer radius of 1.15 m. [ATL97].*

Pixels The Pixel Detector precisely measures particle tracks close to the interaction point to assist the momentum and impact parameter measurements and allow space-tagging of τ-leptons and b-quarks. The 3 detector layers are built up from sensors, segmented into 46,080 pixels, each of 50×400 μm, to provide a 12 μm RΦ resolution and 66 μm in Z. In total, nearly 80 million channels are contained within the 1.4 m long, 0.5 m diameter detector.

The pixel layout has changed significantly since the TDR [ATL99]. Due to a revised, multi-layer beam pipe design, the radius of the inner B-layer has increased from 43.0 mm to 50.5 mm [Gad00]. The length of the pixels has also increased from 300 to 400 μm. A pixel support tube will allow the pixel detector to be inserted at a later date, but the material has increased by a factor of 1.5 in radiation length [Gar03a]. The detector may initially be configured with only 2 of the 3 layers.

SCT The SemiConductor Tracker provides eight precision measurements via 4 layers of back-to-back silicon microstrip detector modules. The 80 μm strip pitch ensures a 16 μm RΦ resolution for each module and a stereo angle of 40 mrad allows 580 μm in Z. The detector contains 61 m^2 of silicon detectors, with 6.2 million channels, read out by radiation hard optical links at a total data transfer rate of up to 40 Gbytes s^{-1}. The detector radial coverage of 300-520 mm provides a large lever arm to contribute to track momentum and impact parameter measurements, and aids pattern

recognition: both current track finding algorithms start in the precision trackers (Pixels + SCT).

The barrel region comprises four concentric carbon-fibre cylinders onto which the modules are mounted via brackets. The close placement of the modules requires precise, robotic mounting techniques. Both end-caps contain nine coaxial wheels, each with up to three annular rings of tapered modules. The SCT support structures are complete and are currently being populated with services and modules, as shown in Figure 1.6.

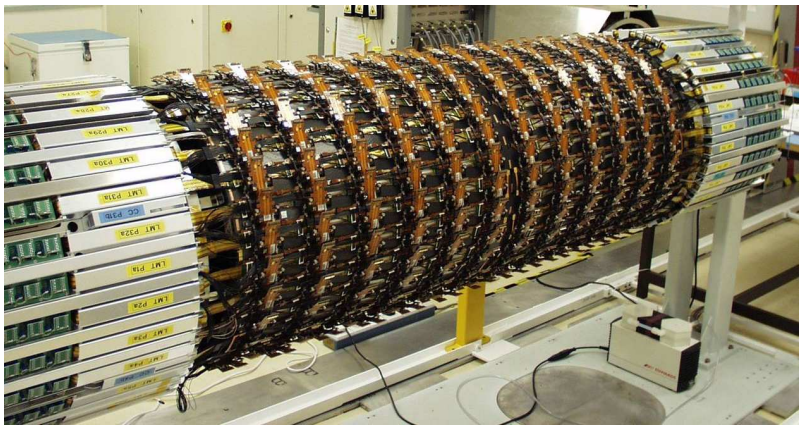

Figure 1.6: *The smallest SCT barrel (1.5 m, Ø 600 mm), complete with readout harnesses, but prior to module mounting.*

TRT The Transition Radiation Tracker provides continuous tracking via an array of Ø 4 mm straw drift-tubes and assists particle identification. A double-threshold readout distinguishes between charge liberated by minimally ionizing particles and signals from transition-radiation produced in polypropylene foils or fibres between the straws. When an ultra-relativistic particle traverses the dielectric boundary, soft X-rays will be produced, which should be detected by the predominantly xenon filled drift-tubes. This enables discrimination of electrons from heavier particle tracks, which is particularly useful in $B_d^0 \rightarrow J/\psi K_s^0$ with $J/\psi \rightarrow e^+e^-$, to overcome the combinatorial background. The 370,000 straws improve the combined detector momentum measurement and pattern recognition and aid the level 2 trigger. Test beam results indicate a $\sim$130 μm RΦ resolution is feasible. The barrel modules are now complete and are being assembled onto the support structures. [Mits03]

1.3 Alignment

The positions of the active detector elements within a particle tracker must be very well known to accurately reconstruct tracks left by long lived charged particles. The procedure of determining the detector positions is called *alignment*. In previous high energy physics experiments, sufficiently precise alignment has traditionally been achieved with track data alone, by examining the hit residuals from the fitted particle trajectories. Aligning the ATLAS Inner Detector solely by this method presents several problems:

- The ATLAS Inner Detector is comparatively enormous: the pixels and SCT are a factor of ten in volume larger than the current largest silicon system (of CDF II). A critical concern for ATLAS is whether the large but low mass tracking detector will be sufficiently stable to permit alignment with tracks. Short-term detector instabilities threaten to severely degrade the alignment precision achievable with tracking data alone (see below).

- The LHC presents new challenges to track based alignment methods, which did not feature at LEP. The well determined beam energies at LEP were entirely transferred via e^+e^- annihilation to the collision debris, providing a valuable alignment constraint. At the LHC, only an *à priori* unknown fraction of the proton's momentum will be carried by the quark or gluon constituents that collide. In addition, the transverse momentum of Z^0 bosons will not be essentially zero as at LEP, so valuable back-to-back dilepton decays cannot be exploited.

- In general, the detector occupancy will be extremely high due to pile-up of minimum bias events. These low momentum events do not aid alignment, due to the dominance of multiple scattering.

- The solenoidal magnetic field will be significantly non-uniform, especially in the forward regions.

- Track based methods alone cannot remove a number of degenerate and weakly constrained modes of detector deformations. [Hin04] (see Chapter 5).

Nevertheless, the high LHC luminosity implies that only twenty four hours of track data should suffice to align the Inner Detector to a statistical precision of few microns, *provided that the detector is stable*. The challenge will be to understand the detector stability, sagitta deformations (see Chapter 5) and associated systematics. Several sources of instability could severely degrade track based alignment methods for the ATLAS SCT. For example, the power

dissipation in the front-end electronics will vary with luminosity by between 15-30 kW, leading to large and varying thermal gradients between the front-end chips and cooling pipes. These will induce distortions of the SCT on time scales potentially shorter than the twenty four hours needed to collect sufficient track data for averaging.

Previous and Proposed HEP Alignment Systems

Several other High Energy Physics experiments have employed or proposed systems to monitor the alignment of tracking detectors.

ALEPH The silicon vertex detector was monitored by an optical fibre laser system that projected intense (6 W) laser pulses (50 ns) once per minute onto the silicon detectors for direct acquisition. Temperature correlated short term (several hours) movements of 5-12 μm were observed, with longer term (months) movements[2] of $\sim$20 μm. Periods of degraded b-tagging performance were identified. The system did not permit complex deformation reconstruction but rigid body alignment corrections were implemented. [Sgu99]

DELPHI The microvertex detector employed capacitive displacement sensors to monitor the alignment. Despite noise problems, movements of up to 15 μm radially and 50 μm in Rϕ were measurable. [Cac92]

ZEUS Laser beams and transparent silicon detectors forming straightness monitors parallel to the beam axis, have been installed in the microvertex detector. [Mat01]

Mark II SLC Capacitive displacement sensors in the vertex detector monitored diurnal variations of $\sim$20 μm. [Bre91]

CDF II A RASNIK system has been installed in the upgraded CDF detector, as described in Chapter 6 [Sal98, Gol02].

CMS An alignment system to monitor the silicon tracker has been proposed consisting of collimated laser beams traversing, primarily end-cap modules. The system is designed to provide an absolute accuracy of 100 μm with a relative positional precision of 10 μm. A LINK laser system connects the tracker alignment system with the muon alignment system, which is based on point sources (LEDs) and CCDs. [Ost01]

ALICE An alignment system for the Inner Tracking System is under development, based on microtelescopes and CCDs. [ALI99]

[2]Fibre pointing stability potentially contributed to the apparent long term movement.

An Alignment System for the ATLAS SCT

The challenge of aligning the ATLAS Inner Detector, has stimulated the development of novel optical techniques for in-situ monitoring of particle detectors. An alignment system has been developed [How01] that will measure the SCT shape on a time scale of a few minutes. The alignment system is based on a geodetic grid of length measurements between nodes attached to the SCT. Combining these measurements will allow the node positions to be reconstructed. These node positions will be interpolated to determine the detector module coordinates. Lengths in the geodetic grid will be measured simultaneously using Frequency Scanning Interferometry (FSI). The system and its implementation are detailed in Chapters 2 and 5.

The SCT alignment system is designed to function during the operation of ATLAS, so that runtime alignment corrections can be applied, potentially improving trigger efficiencies. The system will complement track data in the complete alignment strategy, as described in Chapter 5. The main benefits of the ATLAS SCT alignment system are:

- Continuous alignment during operation of ATLAS.
- An understanding of detector shape on day one of physics.
- Direct measurements of short timescale motions that degrade track based alignment.
- Direct measurements of complex distortions that cannot be determined with tracks alone.
- The system will quicken the detector commissioning phase and rapidly enhance the sensitivity of ATLAS to discovery and precision physics, beyond that achievable with tracks alone.

1.4 Conclusion

Within the next decade, the ATLAS experiment at the LHC will probe the nature of electroweak symmetry breaking and explore an energy regime that promises to revolutionize our understanding of the fundamental constituents of matter. The challenging LHC environment places extraordinary demands on detector technologies and on the alignment of suitable tracking detectors. A novel alignment system has been developed that will rapidly enhance the sensitivity of ATLAS to discovery and precision physics. This thesis contributes to these developments.

Chapter 2

Length Measurements with FSI

2.1 Introduction

Frequency Scanning Interferometry (FSI) enables remote, multiple, simultaneous and precise length measurements. This chapter concerns the development of FSI as the basis for the ATLAS SCT alignment system, following the motivation in Chapter 1.

An FSI system has been developed to overcome the challenges of making robust, precise measurements within the hostile, inaccessible environment of the ATLAS SCT, as discussed in Section 2.2.

Section 2.3 introduces the principles of FSI and reviews the prototype FSI system and the FSI data analysis technique that was developed in previous work. Advances in the FSI technique that improved the reliability and precision of the measurements, are detailed in Section 2.4. These advances were required for the FSI grid measurements described in Chapters 3 and 4.

The ATLAS FSI system must be tolerant to spurious reflections and to varying thermal gradients that modulate the refractive index of the gas within the SCT, as discussed in Sections 2.5 and 2.6 respectively. Adverse effects on FSI from poor quality retro-reflectors were investigated, as in Section 2.7. The FSI components have been designed to minimize spurious effects, as summarized in Section 2.8.

2.2 ATLAS SCT Alignment System

2.2.1 Alignment System Concept

An alignment system based on Frequency Scanning Interferometry (FSI) has been developed to monitor shape deformations of the ATLAS semiconductor

tracker (SCT) [How01]. The alignment system consists of a network of interferometers that measure the distances between nodes attached to the SCT. The distance measurements form a geodetic grid, which is solved to reconstruct the node positions. The requirement that grid shape variations are reconstructed to 10 μm in three dimensions, as discussed in Section 5.2.1, imposes a precision of 1 μm on each one dimensional distance measurement. Data from the FSI system is combined with track and initial X-ray survey data and with information from FEA models of the SCT support structure in a coherent alignment strategy.

FSI relies on scanning the frequency of a narrow line-width laser to induce simultaneous phase shifts in multiple interferometers. The length of each interferometer is determined by comparing its phase shift with the phase shift in a reference interferometer. The ratio of phase shifts gives the ratio of interferometer lengths. The technique is a form of absolute distance interferometry: the lengths are determined afresh with each FSI scan, rather than observing length differences by continuously monitoring the interferometer phase. The FSI system can therefore be powered down as necessary, without penalty in subsequent FSI measurements.

The FSI system provides precise, rather than accurate[1], length measurements relative to a thermally stabilized, evacuated reference interferometer. The measurements need to be only precise, rather than accurate and traceable to the internationally defined metre, primarily because the location of the geodetic grid nodes will not be known to micron level accuracy relative to the module co-ordinates. This precludes the possibility of an absolute calibration of the SCT alignment system, even with accurate length measurements. Instead, the FSI system should be calibrated by an X-ray survey [Dou01] of the SCT that determines the initial positions of all silicon modules. During ATLAS operation, consecutive FSI scans will precisely follow short timescale movements of the silicon modules induced by deformations of the SCT structure, allowing appropriate corrections to be applied to track based alignment methods. The calibration and implementation of the FSI system in ATLAS is discussed further in Section 5.5.

2.2.2 Grid Line Interferometer

The hostile, inaccessible environment of the ATLAS SCT imposes strict requirements on the design of the interferometers that constitute the on-detector FSI system. The interferometers should:

- be remotely measured.
- be tolerant to harsh radiation. Particle collisions at the centre of ATLAS will subject the SCT to a gamma dose of 10 MRad (in silicon) over 10

[1]Precision represents the spread of measurements about the mean, whereas accuracy qualifies the relation of the mean to the true value.

years and a total flux of around 10^{14} (1 MeV equivalent) neutrons [ATL97, Tur01].

- have components of minimal mass and radiation length to reduce multiple scattering of elementary particles.
- have compact components that operate within the confined space between SCT modules, services and support structures.
- continue to function for the $\gtrsim$10 year lifetime of ATLAS without access for adjustment or maintenance.

The solution is a fibre coupled interferometer for each line in the ATLAS grid. These *Grid Line Interferometers* (GLI) consist of two components with a compact, rugged, low mass, zero-maintenance design, as shown in Figure 2.1. The main component is called a *quill* consisting of two parallel single-mode fibres and a fused silica beam-splitter. The other component is an external retro-reflector. The GLI components are clustered into *jewels*[2] that are secured to the SCT support structure to form the grid nodes. A quill in one jewel points to a retro-reflector in another jewel to form a GLI. The distance between the quill and retro-reflector is measured by FSI.

[2]FSI jewels are described in Section 3.3.2.

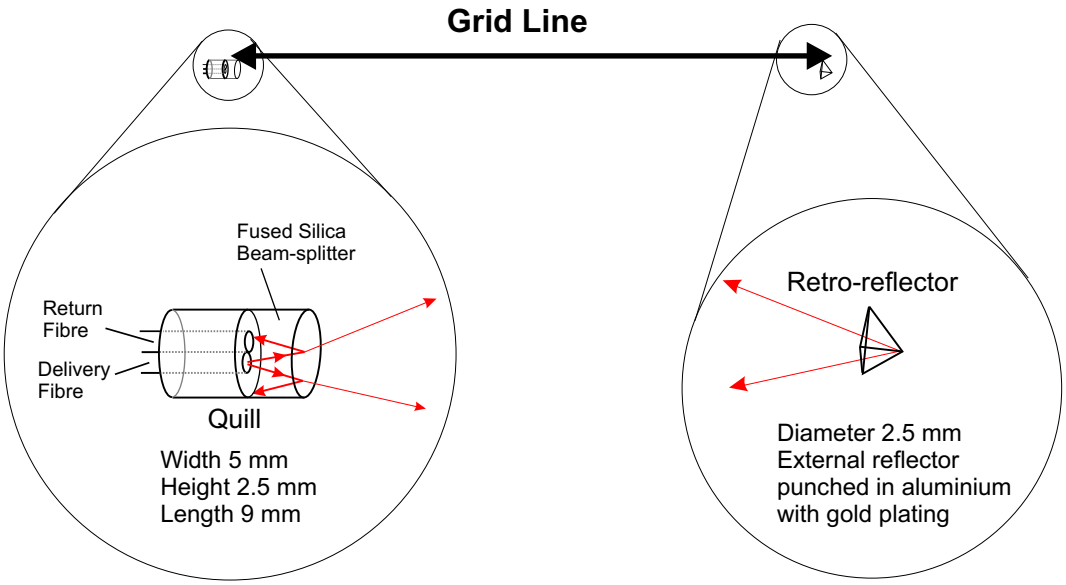

Figure 2.1: *The Grid Line Interferometer design for the ATLAS FSI system. The wide angle beam emerging from the quill provides tolerance to small misalignments which may occur during the 10 year operational lifetime. As a trade off, the interferometer provides a return signal of around only 1 pW per mW of input power, for a 1 m GLI length.*

Light enters the interferometer via the delivery fibre. A partial reflection from the beam-splitter forms the short path of the interferometer (the V-mode), which is combined with a long path via the retro-reflector (the R-mode). The resulting interference, described in Equation 2.4 is coupled into the return fibre to a remote avalanche photodiode (APD).

The plane wave approximations at the return fibre for the spatial complex amplitude of the two reflection modes are,

$$u_{\mathrm{V,R}} = U_{\mathrm{V,R}} e^{i\theta_{\mathrm{V,R}}}. \tag{2.1}$$

The measured quantity is the intensity, which by the superposition principle of light waves is the absolute square of the sum of these complex amplitudes [Hec87]:

$$\begin{aligned} I = |u_{\mathrm{V}} + u_{\mathrm{R}}|^2 &= |u_{\mathrm{V}}|^2 + |u_{\mathrm{R}}|^2 + u_{\mathrm{V}} u_{\mathrm{R}}{}^* + u_{\mathrm{V}}{}^* u_{\mathrm{R}} & (2.2)\\ &= U_{\mathrm{V}}{}^2 + U_{\mathrm{R}}{}^2 + 2\, U_{\mathrm{V}} U_{\mathrm{R}} \cos(\theta_{\mathrm{V}} - \theta_{\mathrm{R}}). & (2.3) \end{aligned}$$

If $P_{\mathrm{V,R}}$ denote the powers of the reflection modes coupled into the return fibre, then the power of the resulting interference is,

$$P_{\mathrm{FSI}} = P_{\mathrm{V}} + P_{\mathrm{R}} + 2\sqrt{P_{\mathrm{V}} P_{\mathrm{R}}}\, \cos(\Theta_{\mathrm{GLI}}) \tag{2.4}$$

for the phase difference,

$$\Theta_{\mathrm{GLI}} = \frac{2\pi}{c} \nu \mathcal{D} \tag{2.5}$$

where,
ν = the frequency,
$\mathcal{D}$ = the optical path difference of the GLI.

Scanning the optical frequency therefore creates sinusoidal variation in the interference signal in the return fibre that is measured by the APD. The measured mean power, P_{DC} and mean amplitude, P_{AC}, of the FSI fringes are:

$$\begin{aligned} P_{\mathrm{DC}} &= P_{\mathrm{R}} + P_{\mathrm{V}} & (2.6)\\ P_{\mathrm{AC}} &= 2\sqrt{P_{\mathrm{R}} P_{\mathrm{V}}}. & (2.7) \end{aligned}$$

Note that[3], assuming $P_{\mathrm{V}} > P_{\mathrm{R}}$,

$$\frac{P_{\mathrm{V}}}{P_{\mathrm{R}}} = \frac{1 + \sqrt{1-\gamma^2}}{1 - \sqrt{1-\gamma^2}} \tag{2.8}$$

where the visibility, $\gamma = \frac{P_{\mathrm{AC}}}{P_{\mathrm{DC}}}$.

The beam-splitter design was tuned to maximize the signal to noise and suppress unwanted multiple reflections (termed W-mode) within the beam-splitter [Coe01, Mit02]. The beam-splitters installed in ATLAS have a thickness of 2.5 mm and the front surface angled by 4.0° [Qui02].

[3]Equation 2.8 is used in Chapter 3

2.3 Frequency Scanning Interferometry

This section reviews the culmination of previous developments of the FSI technique for the ATLAS SCT alignment system. [Nic96, Fox96, Gre99, Coe01, Mit02]

2.3.1 Basic Principle of FSI

FSI is a technique for multiple, simultaneous and precise distance measurements. A narrow line-width tunable laser simultaneously illuminates multiple interferometers (the GLIs) to be measured and a reference interferometer. As the optical frequency is scanned, a phase shift is induced in all interferometers, at a rate that is proportional to the optical path difference (OPD) of each interferometer. From Equation 2.5 and assuming a fixed OPD, the phase shift induced in a GLI is,

$$\Delta\Theta_{\mathrm{GLI}} = \frac{2\pi}{c}\Delta\nu\mathcal{D} \tag{2.9}$$

where,
$\Delta\nu$ = the frequency scan interval,
$\mathcal{D}$ = the OPD of the GLI.
By analogy, the simultaneously induced phase shift in the reference interferometer is,

$$\Delta\Phi_{\mathrm{REF}} = \frac{2\pi}{c}\Delta\nu\mathcal{L} \tag{2.10}$$

where $\mathcal{L}$ = the OPD of the reference interferometer.

The phase shifts in the interferometers are compared to determine the ratio of OPDs:

$$q = \frac{\mathcal{L}}{\mathcal{D}} = \frac{\Delta\Phi_{\mathrm{REF}}}{\Delta\Theta_{\mathrm{GLI}}} \tag{2.11}$$

The length[4] of any GLI, D_{GLI}, is determined from this phase ratio, q, the calibrated length, L, of the reference interferometer and the refractive indices $n_{\mathrm{REF,GLI}}$, of the gases within the interferometers:

$$D_{\mathrm{GLI}} = \frac{L}{q}\frac{n_{\mathrm{REF}}}{n_{\mathrm{GLI}}}. \tag{2.12}$$

In practice, any drift in OPD of an interferometer induces a large error on the phase ratio estimate, q, for short FSI scans. These errors are significantly reduced by tuning the frequency of two lasers in opposite directions, so that

[4]The physical path difference between the two interferometer arms.

the drift errors cancel to first order, as explained in Section 2.3.3. In isolation, the dual laser drift cancellation technique does not provide the required precision for the ATLAS FSI system. The required precision is achieved by drastically extending the effective interval of the frequency scan, as described in Section 2.3.3.

2.3.2 Prototype FSI System

Optical Layout

The FSI demonstration system, developed in previous work, was connected to the prototype grid testing apparatus, described in Chapter 3 to form the prototype FSI system, shown in Figure 2.2. The prototype grids were evaluated in this system, as described in Chapter 4.

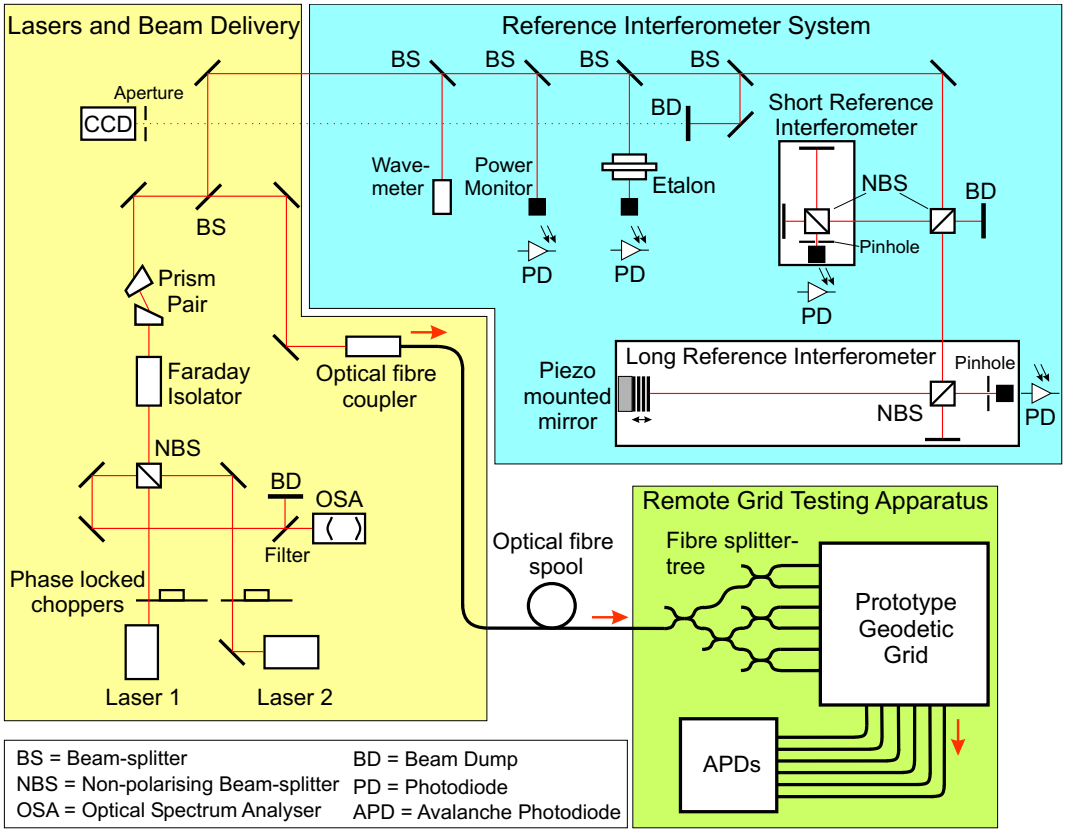

Figure 2.2: *The optical layout of the prototype FSI system comprising 3 parts: the dual laser and beam delivery system, the reference interferometer system, and the remote grid testing apparatus described in Chapter 3.*

The dual laser and reference interferometer system was arranged on a steel optical table that was isolated from extraneous vibrations by pneumatic dampers. The function of each optical component is summarized below. Further details are documented elsewhere [Coe01, Mit02].

The following components formed the dual laser and beam delivery system:

Lasers Two narrow line-width (<300 kHz) coherent light sources were provided by a pair of semiconductor tunable lasers that shared an external cavity diode Littman-Mettcalf design [Har91]. The cavity length could be modulated via a piezo-electric actuated mirror to tune each laser frequency; over a range of 67 GHz for Laser 1 (New Focus 6226) and 93 GHz for Laser 2 (New Focus 6300). The wavelength of Laser 1 could also be set in the range 818 nm to 851 nm in 0.01 nm steps. This enabled the coarse tuning required to drastically extend the frequency interval of the FSI scan, as in Section 2.3.3. Laser 2 had a midpoint wavelength of 836 nm.

Choppers (New Focus 3501) A pair of choppers were phase locked and arranged so that only one laser illuminated the interferometers during any data acquisition cycle. In some FSI scans, extreme jitter of the chopper phase led to synchronization failure, called chopper cross-over, that caused severe glitches in the unwrapped long reference phase, as in Section 2.3.3.

Optical Spectrum Analyser Comprising a scanning Fabry-Pérot confocal etalon and photodiode, this device enabled the laser spectra to be monitored during tuning for phase continuity and single-mode operation.

Faraday Isolator The lasers were isolated from back reflections via downstream components, to guard against tuning disruption. These reflections were attenuated by $\gtrsim$30 dB with this device, which exploited the Faraday effect to rotate by 45° the linear polarization axes of through going light. Back reflections underwent a total rotation of 90° and were blocked by a polarizer nearest the lasers.

Anamorphic Prism Pair The elliptical profiles of the laser beams were approximately circularized to improve fibre coupling and interferometer visibility.

Optical Fibre Coupler Light was focussed and launched into the fibre that conveyed power to the prototype geodetic grid. The input end of the fibre was finely positioned using a 3 axis stage, to coincide with the focus of the incoming laser beam.

CCD A camera was added to the system to aid realignment of Laser 2, which had been removed and repaired. While Laser 1 illuminated the system, the fibre coupler was adjusted to maximize the power coupled into the fibre. The beam of Laser 2 was then iteratively steered with two mirrors to maximize the light coupled from Laser 2 into the fibre and simultaneously ensure both laser spots were coincident on the CCD. This procedure optimized the collinearity of the laser beams. The visibilities in the reference interferometers were checked and a small improvement ($\frac{\Delta\gamma_{\mathrm{LR}}}{\langle\gamma_{\mathrm{LR}}\rangle} \sim 13\%$, $\frac{\Delta\gamma_{\mathrm{SR}}}{\langle\gamma_{\mathrm{SR}}\rangle} \sim 2\%$) was found compared to the visibilities for the alignment of Laser 2 before repair. During FSI measurements, the optical path to the CCD camera was blocked.

The remaining components formed the reference interferometer system, that precisely measured the frequency tuning interval:

Wave-meter (Burleigh 2100) A scanning Michelson interferometer provided a check of the wavelength of each laser to $\sim\pm0.1$ nm.

Power Monitor A photodiode (Thorlabs PDA50) monitored variations in optical power entering the reference interferometer system.

Etalon The coarse tuning frequency interval of Laser 1 was monitored by an air spaced Fabry-Pérot etalon (Tec Optics EA 202) that had a free spectral range of 10 GHz.

Long Reference Interferometer (LR) This defined the reference length against which all other interferometers were compared. Stability of this length was paramount to the repeatability of FSI measurements. The design of this Michelson interferometer exploited low CTE ($\sim$1 ppm K^{-1}) invar alloy rods to reduce thermally induced length changes. The LR was surrounded by thermal insulation and one end was supported on Teflon sheets to reduce friction caused by differential thermal expansion of the optical bench relative to the LR. The physical path difference between the two interferometer arms was assumed fixed at the previously estimated value of 859.09 mm [Coe01].

Piezo mounted mirror The mirror in the long arm of the LR was mounted on piezo-electric transducer. The mirror was displaced to modulate the OPD and implement phase stepping, as described in Section 2.3.2.

Short Reference Interferometer (SR) This Michelson interferometer shared the LR design, but had a shorter length of $\sim$170.42 mm. The SR was essentially a stable reference GLI, used to resolve ambiguities in the

data analysis by comparing the measured SR length with its previously calibrated length. This verification process was automated, as described in Section 2.4.3.

The grid testing apparatus is detailed in Chapter 3.

Data Acquisition (DAQ) System

A functional overview of the hardware control and DAQ scheme is shown in Figure 2.3.

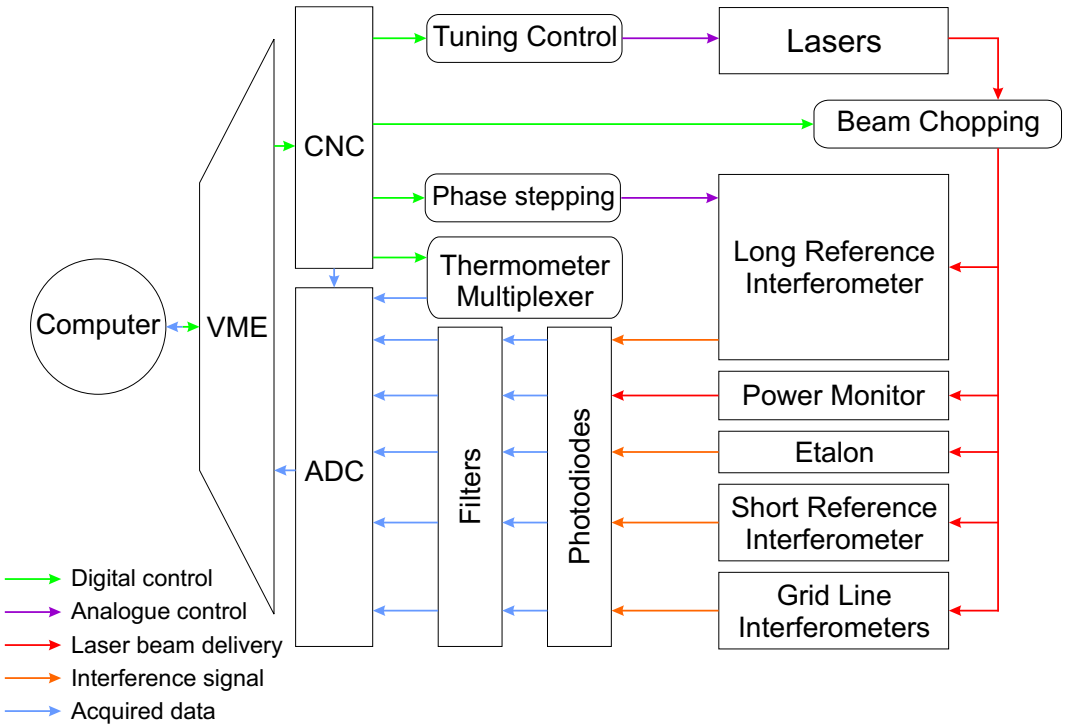

Figure 2.3: *Control and data acquisition for the prototype FSI system.*

A computer controlled the hardware and read out the acquired data in real time via a National Instruments VME-MXI-2 interface to two cards:

Clock and Control Card (CNC) [**Mit02**] A custom-built card generated synchronized control signals for laser tuning, beam chopping, phase stepping and thermometer multiplexing.

Burr-Brown MPV 901 ADC The filtered analogue signals from the photodiodes and the multiplexed thermometer signals were digitized by this

card and read out via VME to the computer. The 12 bit ADC had 16 channels.

At the start of an FSI scan the computer also directly set via GPIB[5] the start and stop wavelengths that defined the total coarse tuning frequency interval of Laser 1.

DAQ Cycle

Data was acquired in two operational modes defined by the rate of laser frequency tuning: *fine tuning* and *coarse tuning*. The devices recorded in each operational mode are shown in Table 2.1. The read out sequence formed one *DAQ cycle* which repeated every 2 ms. All devices were recorded during fine tuning but only the etalon and the synchronization signal were recorded during coarse tuning. This quasi-continuous monitoring of the etalon was required to acquire sufficient data to count the narrow resonant frequency etalon peaks in an automated procedure, for the purposes described in Sections 2.3.3 and 2.4.4.

DAQ sequence	Fine Tuning	Coarse Tuning
0	LR	Etalon
1	Etalon	Etalon
2	SR	Etalon
3	APD 0	Etalon
4	APD 1	Etalon
5	APD 2	Etalon
6	APD 3	Etalon
7	APD 4	Etalon
8	APD 5	Etalon
9	APD 6	Etalon
10	SYNC	SYNC
11	Power Monitor	Etalon
12	Thermometer Multiplexer	Etalon

Table 2.1: *The sequence of a DAQ cycle during fine or coarse tuning. The cycle repeated every 2 ms for the clock rate set.*

LR Phase Stepping

The choppers were synchronized so that each laser alternately illuminated the reference interferometer system for 4 consecutive DAQ cycles, to allow the

[5]General Purpose Interface Bus

LR phase to be measured. During these 4 cycles the piezo mounted mirror of the LR was displaced to modulate the LR OPD. The piezo displacement was approximately linear and the time intervals between 4 LR intensity measurements were approximately equal. The measurement were therefore made at equal steps in LR phase, necessary for the phase extraction algorithm described in Section 2.3.3.

One extra redundant DAQ cycle was allowed between alternate groups of 4 DAQ cycles, to contain chopper cross-over between lasers. The chopper cycle therefore repeated every 10 DAQ cycles as shown in Figure 2.4. The DAQ clock rate from previous work was retained at 500 Hz so the chopper cycle rate was 50 Hz.

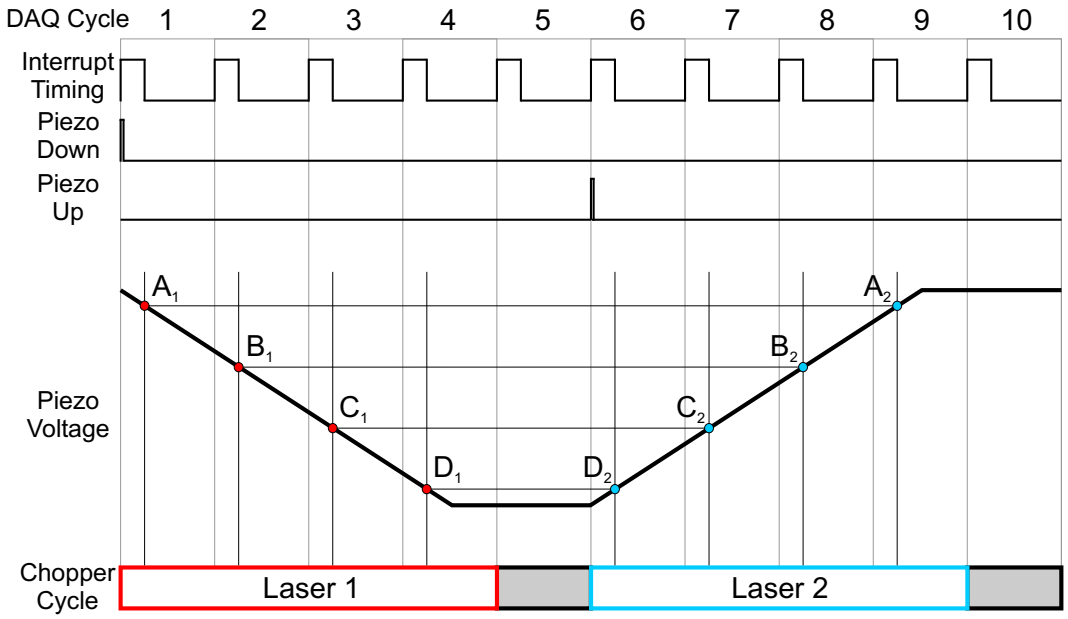

Figure 2.4: *The sequence of 10 DAQ cycles in a chopper cycle. The voltage ramp that controlled the LR piezo to implement phase stepping is shown. This chopper cycle repeated every 20 ms for a DAQ rate of 500 Hz.*

The CNC card controlled the timing of the voltage ramp that set the LR piezo in motion. These timing signals were adjusted so that the 4 positions of the piezo when the LR intensity was read, were equal for both lasers. Any mismatch would lead to a systematic error between lasers during LR phase extraction. A *piezo balancing* procedure similar to that documented elsewhere [Mit02] was followed to determine the relevant pulse width shown in Table 2.2.

Control Pulse	Pulse width [μs]
Interrupt Timing	511
Piezo Up	50
Piezo Down	50

Table 2.2: *The width of the timing pulses determined by the piezo balancing procedure, for a DAQ rate of 500 Hz.*

2.3.3 FSI Data Analysis

The techniques described in this section were developed in earlier work [Coe01, Mit02]. Advances on this previous work are described in Section 2.4.

Overview

An FSI measurement was performed by monitoring the phase shift in all interferometers as the optical frequency was scanned. Continuous changes in phase of the LR could only be directly measured during fine tuning. In isolation, the fine tuning ranges of the lasers were insufficient to achieve the required precision. Therefore Laser 1 was coarse tuned between fine tuning *subscans* to drastically extend the frequency interval of the complete FSI scan, as shown in Figure 2.5. Laser 2 could not be coarse tuned. During fine tuning subscans Laser 2 was tuned in the opposite direction to Laser 1, to implement the drift

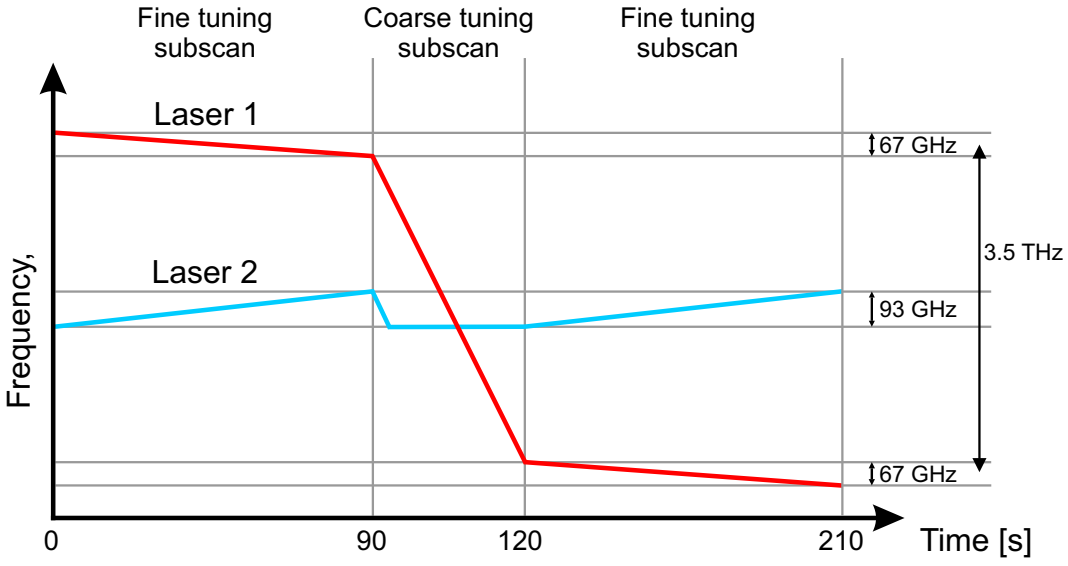

Figure 2.5: *An overview of the frequency tuning subscan sequence that formed a basic FSI scan. The fine tuning ranges of both lasers are exaggerated.*

cancellation technique described below. During coarse tuning, Laser 2 was simply rewound ready for the next fine tuning subscan. The etalon monitored the coarse tuning interval of Laser 1 so that the total phase shift could be determined by *linking* the fine tuning subscans, as described below.

LR Phase Extraction and Unwrapping

The change in laser frequency was monitored during fine tuning by measuring the phase change in the LR. The LR phase was extracted by a modified Carré algorithm [Coe01] detailed below, from the 4 LR intensity measurements recorded for each laser during a chopper cycle. The phase steps were assumed to be of equal size, α_S, with $0 < \alpha_S < \pi$, and centred on the phase Φ_{MID}. The LR intensity for each phase step, n=1...4, was,

$$I_n = I_0 \left[1 + \gamma \cos \left(\Phi_{MID} + \left(\frac{5}{2} - n \right) \alpha_S \right) \right] \tag{2.13}$$

where γ =visibility and I_0=mean LR intensity. The phase step size was given by,

$$\alpha_S = \arccos \left[\frac{1}{2} \left(\frac{I_1 - I_4}{I_2 - I_3} - 1 \right) \right] \tag{2.14}$$

provided $(I_2 - I_3) \neq 0$, otherwise:

$$\alpha_S = \begin{cases} 0 & \text{if} \quad (I_2 > I_1) \text{ AND } (I_3 > I_4) \\ \pi & \text{if} \quad (I_2 < I_1) \text{ AND } (I_3 < I_4) \end{cases} \tag{2.15}$$

or exceptionally, an error had occurred. The extracted LR phase modulo π, called the *local phase*,

$$\phi = \arctan \left[\frac{\tan(\frac{1}{2}\alpha_S)(|I_2 - I_4| + |I_1 - I_3|)}{|I_2 - I_4| - |I_1 - I_3|} \right] \tag{2.16}$$

The interferometer phase Φ is the sum of the local phase ϕ and a multiple m of 2π. The following notation is introduced [Coe01]:

$$\begin{aligned} \Phi &= \phi + 2m\pi \\ &= \langle \phi, m \rangle \end{aligned} \tag{2.17}$$

where $-\pi \leq \phi < \pi$ and m is an integer, called the *order number*. The operator **fold**() reduces the phase Φ to the local phase:

$$\textbf{fold}(\Phi) = \phi \tag{2.18}$$

The folded, extracted local LR phases ϕ_i from each chopper cycle, i, of a fine tuning subscan that survived extraction errors [Coe01] were unwrapped with the aim of producing a continuous function of LR phase measurements, Φ_i. The original phase unwrapping and glitch finding algorithms [Coe01] did not correct all phase discontinuities that disrupted FSI measurements. The phase discontinuities arose from periods of multi-moded lasing or chopper cross-over. The increasing propensity of multi-moded lasing motivated the development of a new phase unwrapping algorithm, described in Section 2.4.2.

GLI Intensity Sine Fitting

The unwrapped LR phase provided a frequency axis against which the other interferometer phases were compared. No phase stepping was performed in the other interferometers. The required LR to GLI phase ratio, q, from Equation 2.11, was estimated by fitting the intensity of the GLI interference fringes versus the LR phase, with,

$$P_{\mathrm{FSI}} = P_{\mathrm{DC}} + P_{\mathrm{AC}} \cos(\frac{\Phi_i}{q} + \alpha) \tag{2.19}$$

from Equations 2.4, 2.6 and 2.7 and where α is an arbitrary phase offset. MINUIT [Jam75] was used to minimize the fit quality parameter,

$$\chi^2 = \sum_i \frac{[P_{\mathrm{FSI}}(P_{\mathrm{DC}}, P_{\mathrm{AC}}, q, \alpha, \Phi_i) - Y_i(\Phi_i)]^2}{P_{\mathrm{AC}}} \tag{2.20}$$

where the summation is over all synchronous data points in a fine tuning subscan and $Y_i(\Phi_i)$ is the GLI intensity data.

For perfectly sinusoidal interference the variation of χ^2 with q takes the form of an inverted sinc function [Coe01], exhibiting a distinct, deep minimum into which the fit should settle. Deviations from this form indicate an error, such as a phase glitch, or poor quality GLI fringes, that typically resulted in a large systematic error. This χ^2 *profile* was therefore a useful diagnostic tool.

Sine fits were performed for all interferometer data from all fine tuning subscans and for both lasers. The sine fitting dominated the computation time of the analysis. For 12 interferometers[6] and 4 fine tuning subscans the sine fitting took ∼2.3 minutes on a 1.6 GHz Pentium® 4 processor. A small PC farm is envisaged to process data for the 842 GLIs of the ATLAS FSI system, to provide the SCT alignment corrections in quasi real time.

[6] 7 GLIs, SR and 4 pseudo-GLIs from the 4 phase steps of the LR for cross-checking.

Dual Laser Drift Cancellation

If the optical path difference of an interferometer linearly drifts during an FSI scan, then the phase shift equations, 2.9 and 2.10, are modified to:

$$\Delta\Theta_{\mathrm{GLI}} = \left(\frac{2\pi}{c}\right)(\mathcal{D}\Delta\nu + \nu\Delta\mathcal{D} + \Delta\nu\Delta\mathcal{D}) \tag{2.21}$$

$$\Delta\Phi_{\mathrm{REF}} = \left(\frac{2\pi}{c}\right)(\mathcal{L}\Delta\nu + \nu\Delta\mathcal{L} + \Delta\nu\Delta\mathcal{L}) \tag{2.22}$$

If $\frac{\Delta\nu}{\nu} \ll 1$ higher order terms may be neglected, and Equation 2.11 is modified to,

$$q = \frac{\Delta\Phi_{\mathrm{REF}}}{\Delta\Theta_{\mathrm{GLI}}} \tag{2.23}$$

$$= \frac{(\frac{2\pi}{c})\left[\mathcal{L}\Delta\nu + \nu\Delta\mathcal{L}\right]}{(\frac{2\pi}{c})[\mathcal{D}\Delta\nu + \nu\Delta\mathcal{D}]} \tag{2.24}$$

$$= \frac{\mathcal{L}\left[1 + (\frac{\nu}{\Delta\nu})(\frac{\Delta\mathcal{L}}{\mathcal{L}})\right]}{\mathcal{D}[1 + (\frac{\nu}{\Delta\nu})(\frac{\Delta\mathcal{D}}{\mathcal{D}})]}. \tag{2.25}$$

If $\frac{\Delta\mathcal{L}}{\mathcal{L}}, \frac{\Delta\mathcal{D}}{\mathcal{D}} \ll \frac{\Delta\nu}{\nu}$ then by a Taylor expansion of the denominator,

$$q \approx \left(\frac{\mathcal{L}}{\mathcal{D}}\right)(1 + \Omega\varepsilon) \tag{2.26}$$

where,

$$\Omega = \frac{\nu}{\Delta\nu} \quad , \quad \varepsilon = \frac{\Delta\mathcal{L}}{\mathcal{L}} - \frac{\Delta\mathcal{D}}{\mathcal{D}}. \tag{2.27}$$

Any relative interferometer drift error, ε, is therefore magnified by a factor Ω. In the prototype FSI system the fine tuning range $\Delta\nu \approx 70$ GHz and the mean laser wavelength, $\lambda \approx 836$ nm, so $\Omega \approx 5000$. Therefore a 1 nm relative linear drift between interferometers of length 1 m induces a 5 ppm error (equivalent to 5 μm length error) in the phase ratio q_k, determined from the above GLI intensity sine fitting for a single laser, $k = 1, 2$, in a single fine tuning subscan:

$$q_k \approx \left(\frac{\mathcal{L}}{\mathcal{D}}\right)(1 + \Omega_k\varepsilon_k). \tag{2.28}$$

Importantly, by scanning the laser frequencies in opposite directions, the phase ratios may be combined to cancel the linear drift error.[7]

$$q_0 = \frac{q_2 - \rho q_1}{1 - \rho} = \frac{\mathcal{L}}{\mathcal{D}} \tag{2.29}$$

[7]Opposite direction frequency scanning is not strictly required to reduce drift errors, but tuning the lasers in the same direction was found to provide inadequate drift correction due to increased sensitivity to errors in ρ [Coe04].

where $\rho = \frac{\varepsilon_2 \Omega_2}{\varepsilon_1 \Omega_1}$. The measurements are conducted to ensure the drift errors, ε_1 and ε_2, are approximately equal for both lasers, so that the ratio of drift error terms, ρ, may be decoupled from interferometer drift, $\rho \approx \frac{\Omega_2}{\Omega_1}$.

An example of dual laser drift cancellation is shown in Figure 2.6.

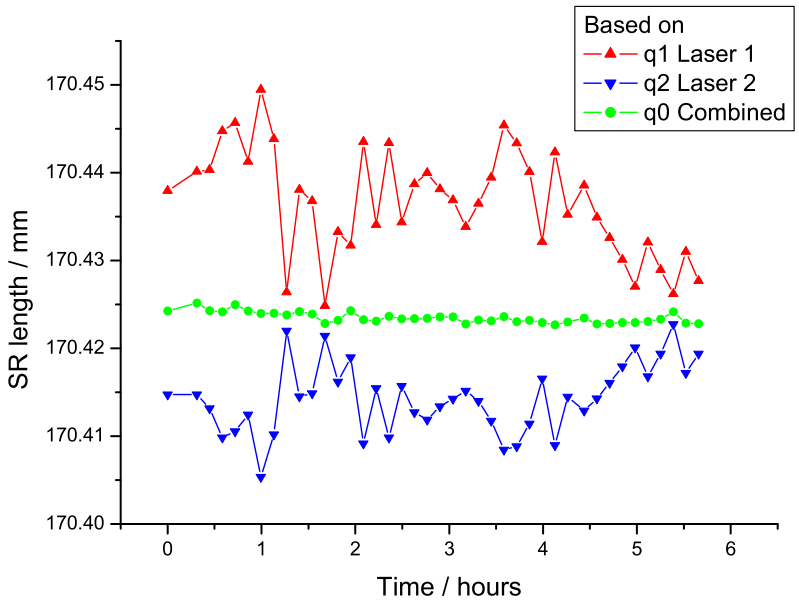

Figure 2.6: *Dual laser drift cancellation: the phase ratio estimate from two lasers scanning in opposite directions may be combined to cancel linear drift errors.*

Subscan Linking

The above *dual laser drift cancellation* technique significantly improves the estimate of the phase ratio, q, corresponding to the ratio of interferometer lengths. Despite this improvement, the 1 μm length measurement precision required for the ATLAS FSI system, is not directly met using the phase ratio estimated from only one fine tuning subscan. The precision is primarily limited by the fine tuning range of the lasers (tens of GHz) and the residual GLI phase errors from the sine fit. The precision can be improved by extending the frequency interval of the scan, which importantly reduces the relative contribution of the GLI phase errors to the ratio of interferometer lengths.

In practice, the frequency interval is extended by coarse tuning one laser between two fine tuning subscans. The coarse tuning provides a large and rapid shift in frequency over several THz. The advantage of this method is that the large frequency interval is spanned rapidly, to minimize the total scan

time and thus reduce susceptibility to additional interferometer drift errors. The challenge is that the frequency shift during coarse tuning is too rapid for the LR phase to be reliably extracted by the phase stepping method described above. Therefore, the phase difference between the two fine tuning subscans is initially unknown. The process of *linking* the phase information between fine tuning subscans is described in this section. Once the phase has been *linked*, the improved phase ratio is determined from the measured phase difference between the midpoints of the two subscans, labelled A and B:

$$q_{\mathrm{AB}}^{\mathrm{LINK}} = \frac{\Phi_{\mathrm{MID(B)}} - \Phi_{\mathrm{MID(A)}}}{\Theta_{\mathrm{MID(B)}} - \Theta_{\mathrm{MID(A)}}} = \frac{\Delta\Phi_{\mathrm{AB}}}{\Delta\Theta_{\mathrm{AB}}} \tag{2.30}$$

First the LR phase is *linked* from subscan A to subscan B using the change in laser frequency monitored by the etalon to determine $\Delta\Phi_{\mathrm{AB}}$. Only after the LR phase is linked, can the GLI phase also be linked to determine $\Delta\Theta_{\mathrm{AB}}$.

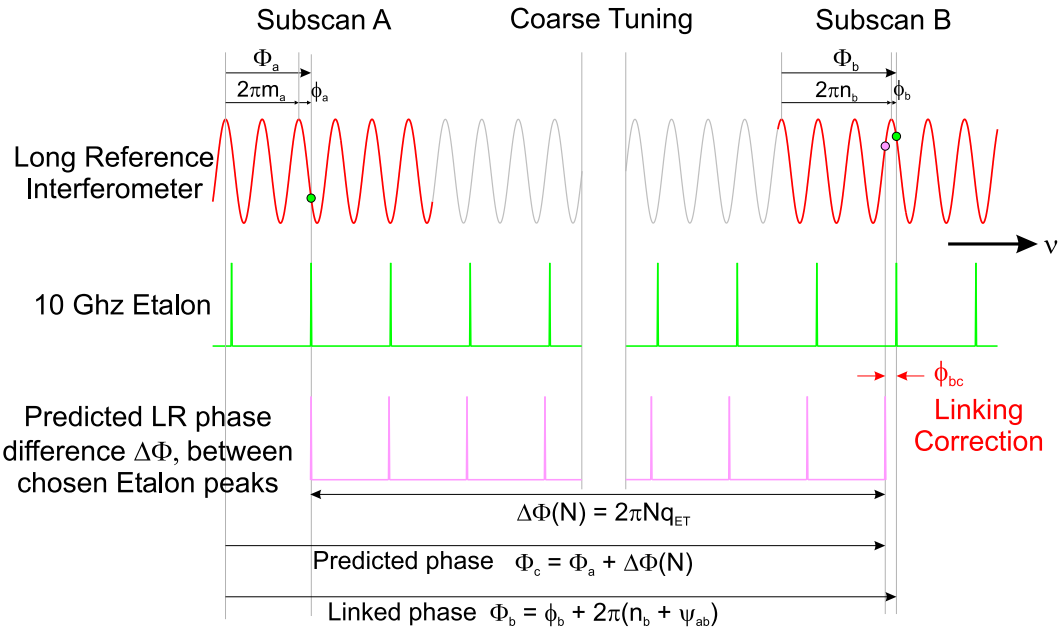

Figure 2.7: *Subscan linking for the Long Reference Interferometer (LR). A comb of etalon peaks create reference markers that enable the phase information from subscan A to be linked to subscan B. The LR phase values at a peak in each subscan are well measured locally (green points), but the phase difference between these points is initially unknown. An extrapolation from the peak in subscan A predicts the phase at a peak in subscan B (pink point). The extrapolation is then corrected using the difference between the predicted phase and the locally measured phase in subscan B.*

The LR phase difference between subscans is determined with the aid of the etalon, which generates a reference comb of peaks with equal frequency spacing, as shown in Figure 2.7. Prior knowledge of the etalon to LR phase ratio allows an extrapolation of the LR phase information from subscan A to subscan B. The locally measured phase in subscan B can be used to correct the prediction, provided that the extrapolation is correct to within $\pm\pi$ as detailed below.

The LR phase difference between the two fine tuning subscans, $\Delta\Phi$, is initially estimated by counting the number of etalon peaks during coarse tuning:

$$\Delta\Phi(\mathrm{N}) = 2\pi \mathrm{N} \mathrm{q}_{\mathrm{ET}} \tag{2.31}$$

where q_{ET} is the best estimate of the phase ratio of the etalon and N is the number of etalon peaks between a peak in subscan A and a peak in subscan B.

This estimated phase difference is added to the LR phase at the peak in subscan A, Φ_a, to provide a prediction, Φ_c of the LR phase at the peak in subscan B:

$$\Phi_c = \Phi_a + \Delta\Phi(\mathrm{N}) = \langle \phi_{\mathrm{c}}, \mathrm{m_c} \rangle \tag{2.32}$$

following the notation introduced on page 25. The difference between the prediction and the locally measured LR phase at the peak in subscan B, Φ_b, provides a correction, (ϕ_{bc}), to the total phase difference:

$$\Phi_b - \Phi_c = (\phi_b - \phi_c) + 2\pi(n_b + \psi_{ab} - m_c) = \langle \phi_{bc}, 0 \rangle \tag{2.33}$$

where ψ_{ab} is the integer multiple of 2π in LR phase between the phase conventions in subscans A and B [Coe01]. Importantly, for this correction to be valid, the prediction must lie within $\pm\pi$ of the locally measured LR phase at the peak in subscan B. Therefore the sum of order number terms on each side of Equation 2.33 *must* be zero. On the left hand side, the order number terms are within the brackets following 2π *and* in the term[8],

$$\phi_b - \phi_c = \mathbf{fold}(\Phi_b) - \mathbf{fold}(\Phi_c) = \langle \phi_{bc}, g \rangle \tag{2.34}$$

where $-1 \leq g \leq 1$. Equating the order number terms in Equations 2.33 and 2.34:

$$\psi_{ab} = m_c - g - n_b \tag{2.35}$$

Once ψ_{ab} is determined, the link corrected phase difference between the two mid-points in subscans A and B may be derived,

$$\begin{aligned} \Delta\Phi_{\mathrm{AB}} &= \Phi_{\mathrm{MID(B)}} - \Phi_{\mathrm{MID(A)}} \\ &= \langle \phi_{\mathrm{MID(B)}}, (n_{\mathrm{MID(B)}} + \psi_{ab}) \rangle - \langle \phi_{MID(A)}, m_{\mathrm{MID(A)}} \rangle \end{aligned} \tag{2.36}$$

[8]Note in general, $\mathbf{fold}(X - Y) \neq \mathbf{fold}(X) - \mathbf{fold}(Y)$, since $-\pi \leq \mathrm{LHS} < \pi$, whereas $-2\pi < \mathrm{RHS} < 2\pi$.

The role the etalon played in monitoring laser tuning for LR phase linking, is taken up by the linked LR phase for linking the GLI phase. The parameters from the best fit sine function provide an estimate of the GLI phase at the mid-points of either subscan, A or B, for a given LR phase:

$$\Theta_{\mathrm{MID(A,B)}} = \frac{\Phi_{\mathrm{MID(A,B)}}}{q_{\mathrm{A,B}}} + \alpha_{\mathrm{A,B}} \tag{2.37}$$

The difference between a GLI and the LR is that the GLI length drifts[9]. Therefore the mid-point GLI phase in subscan B, Θ_{B}, needs to be corrected for the drift that occurred while the laser tuned between mid-points of each subscan:

$$\widetilde{\Theta}_{\mathrm{B}} = \Theta_{\mathrm{B}} - \Delta\Theta^{\mathrm{DRIFT}} \tag{2.38}$$

where $\Delta\Theta^{\mathrm{DRIFT}}$ is the drift correction, explained below. All terms affected by this drift correction are denoted by a tilde, $\sim$. The steps for linking the drift corrected GLI phase are analogous to the steps for LR phase linking.

The GLI phase shift between subscans, $\Delta\Theta$, is initially estimated from the linked LR phase shift and the current best estimate of the GLI phase ratio[10], q_0^{AV}.

$$\Delta\Theta = \frac{\Delta\Phi_{\mathrm{AB}}}{q_0^{\mathrm{AV}}} \tag{2.39}$$

The estimated phase difference is added to the mid-point GLI phase of subscan A, Θ_{A}, to provide a prediction, Θ_C, for the mid-point GLI phase in subscan B,

$$\Theta_{\mathrm{C}} = \Theta_{\mathrm{A}} + \Delta\Theta = \langle \theta_{\mathrm{C}}, M_{\mathrm{C}} \rangle \tag{2.40}$$

As before, the difference between the prediction and the locally measured (drift corrected) mid-point GLI phase of subscan B, $\widetilde{\Theta}_{\mathrm{B}}$, must lie within $\pm\pi$. By analogy with Equations 2.33 and 2.34,

$$\widetilde{\Theta}_{\mathrm{B}} - \Theta_{\mathrm{C}} = (\widetilde{\theta}_{\mathrm{B}} - \theta_{\mathrm{C}}) + 2\pi(\widetilde{M}_{\mathrm{B}} - M_{\mathrm{C}}) = \langle \widetilde{\theta}_{BC}, 0 \rangle \tag{2.41}$$

$$\text{and} \quad \widetilde{\theta}_{\mathrm{B}} - \theta_{\mathrm{C}} = \mathbf{fold}(\widetilde{\Theta}_{\mathrm{B}}) - \mathbf{fold}(\Theta_{\mathrm{C}}) = \langle \widetilde{\Theta}_{\mathrm{BC}}, G \rangle. \tag{2.42}$$

Substituting Equation 2.42 into Equation 2.41,

$$\widetilde{\Theta}_{\mathrm{B}} - \Theta_{\mathrm{C}} = \langle \widetilde{\Theta}_{\mathrm{BC}}, G \rangle + 2\pi(\widetilde{M}_{\mathrm{B}} - M_{\mathrm{C}}) \tag{2.43}$$

Equating the sum of order number terms to zero links the GLI phase between the subscans,

$$\widetilde{M}_{\mathrm{B}} = M_{\mathrm{C}} - G \tag{2.44}$$

[9] LR drift can be attributed to the GLI, in a first order approximation [Coe01].

[10] Before any subscans are linked, q_0^{AV} is taken as the mean q_0 of all subscans.

so that the link corrected GLI phase shift between the two mid-points of subscans A and B,

$$\begin{aligned} \Delta\Theta_{\mathrm{AB}} &= \Theta_{\mathrm{MID(B)}} - \Theta_{\mathrm{MID(A)}} \\ &= \langle \widetilde{\theta}_{\mathrm{B}}, \widetilde{M}_{\mathrm{B}} \rangle - \Theta_{\mathrm{A}} \end{aligned} \tag{2.45}$$

The GLI drift between subscans was monitored using Laser 2, which was fine tuned over the same range, as in Figure 2.5. The drift correction term, $\Delta\Theta^{\mathrm{DRIFT}}$, mentioned above was calculated by linking the GLI phase of Laser 2. The linking for Laser 2, followed the identical procedure outlined above (omitting the GLI drift correction) for Laser 1, to determine the total Laser 2 GLI phase change between the two midpoints of subscans A and B,

$$\Delta\Theta_2^{\mathrm{TOTAL}} = \Theta_{\mathrm{MID(B),2}} - \Theta_{\mathrm{MID(A),2}} \tag{2.46}$$

This total Laser 2 GLI phase change may also be expressed as the sum of phase changes induced by laser tuning and by GLI drift,

$$\Delta\Theta_2^{\mathrm{TOTAL}} = \Delta\Theta_2^{\mathrm{TUN}} + \Delta\Theta_2^{\mathrm{DRIFT}} \tag{2.47}$$

By analogy with Equation 2.39, the current best estimate of the GLI phase ratio, q_0^{AV}, is combined with the linked LR phase difference to predict the GLI phase shift due to tuning Laser 2,

$$\Delta\Theta_2^{\mathrm{TUN}} = \frac{\Delta\Phi_{\mathrm{AB,2}}}{q_0^{\mathrm{AV}}} \tag{2.48}$$

To minimize the impact of errors in q_0^{AV} on this prediction, Laser 2 is tuned to make the frequency difference between the mid-points of each subscan as small as possible. The drift term for Laser 2 is estimated by,

$$\Delta\Theta_2^{\mathrm{DRIFT}} = \Theta_{\mathrm{MID(B),2}} - \Theta_{\mathrm{MID(A),2}} - \frac{\Delta\Phi_{\mathrm{AB,2}}}{q_0^{\mathrm{AV}}} \tag{2.49}$$

The phase change due to linear drift scales with laser frequency, from Equation 2.21, so the equivalent phase change due to drift for Laser 1 is,

$$\Delta\Theta^{\mathrm{DRIFT}} = \Delta\Theta_2^{\mathrm{DRIFT}} \left(\frac{\nu_1}{\nu_2} \right) \tag{2.50}$$

which is the correction term applied in Equation 2.38. The mean frequencies of the lasers were deliberately set so that $\nu_1 \approx \nu_2$.

Equations 2.30, 2.36 and 2.45 are combined to provide the linked phase ratio,

$$q_{\mathrm{AB}}^{\mathrm{LINK}} = \frac{\Delta\Phi_{\mathrm{AB}}}{\Delta\Theta_{\mathrm{AB}}} = \frac{\langle \phi_{\mathrm{MID(B)}}, (n_{\mathrm{MID(B)}} + \psi_{ab}) \rangle - \langle \phi_{MID(A)}, m_{\mathrm{MID(A)}} \rangle}{\langle \widetilde{\theta}_{\mathrm{B}}, \widetilde{M}_{\mathrm{B}} \rangle - \Theta_{\mathrm{A}}} \tag{2.51}$$

Air Refractive Index Corrections

In the prototype FSI system, the LR and GLIs shared the same laboratory air[11]. The refractivity of dry air at 1000 mbar has the following dependency on the temperature, T (in °C), [Coe01]:

$$\frac{\varrho}{10^6} = 266.62 - 0.882(T - 20.0) \tag{2.52}$$

The differing refractive indices of the air within the GLIs and the LR due to temperature were compensated for when determining each GLI length, from Equation 2.12:

$$D_{\mathrm{GLI}} = \frac{L}{q_{\mathrm{AB}}^{\mathrm{LINK}}} \frac{n_{\mathrm{REF}}}{n_{\mathrm{GLI}}} \approx \frac{L}{q_{\mathrm{AB}}^{\mathrm{LINK}}} \left[1 + \varrho(T_{\mathrm{LR}} - T_{\mathrm{GLI}})\right] \tag{2.53}$$

2.4 Advances in FSI Technique

2.4.1 Introduction

The key to precise FSI measurements is to scan the largest possible frequency interval in the shortest possible time. This reduces error contributions from interferometer drift, as in Equations 2.26 and 2.27, to improve the measurement precision.

In previous work [Coe01], FSI measurements over a frequency interval of 3.5 THz were achieved by rapidly coarse tuning one laser between two fine tuning subscans. The frequency interval was limited by errors in linking the interferometer phase between subscans. As discussed in Section 2.3.3, for the linking corrections to be valid, the extrapolation in phase from any point in one subscan to any point in the second subscan, must lie within $\pm\pi$ of the locally measured phase of the point in the second subscan, from Equations 2.33 and 2.41. The error on the best estimate of the phase ratio, q, limits the frequency interval over which two fine tuning subscans can be validly linked and therefore limits the measurement precision.

To improve the measurement precision, the frequency interval may be extended by sequentially linking a series of fine tuning subscans as follows. After linking the first two subscans, the improved estimate of the phase ratio, $q_{\mathrm{AB}}^{\mathrm{LINK}}$ from Equation 2.51, enables a second extrapolation over a larger frequency interval to another fine tuning subscan. This second link should again improve the estimate of the phase ratio, so that the process may be iterated. In this

[11] In the ATLAS FSI system, the reference interferometer is evacuated and the GLIs are immersed in the dry purge gas of the SCT (N_2).

manner, a series of fine tuning subscans, interspersed with periods of coarse tuning, may be sequentially linked in order of increasing frequency interval, until the required precision is achieved.

This technique of *multiple subscan linking* was implemented for the first time, as described in Section 2.4.4. This was an important development for the ATLAS FSI system because the technique improves not only the FSI measurement precision but also the reliability of measurements by building up large frequency intervals from smaller links, chosen to ensure that phase extrapolation errors are well within the required $\pm\pi$.

All the FSI software developed in previous work was rewritten[12],

- to implement multiple subscan linking.
- to implement reliable etalon peak counting and automatic verification with the SR (see Section 2.4.3).
- to allow multiple GLIs to be recorded and analysed.
- to speed up data saving.[13]
- to avoid any unseen bugs propagating from the previous software.
- to rethink every algorithm, such as the phase unwrapping algorithm discussed in Section 2.4.2.

2.4.2 Improved Phase Unwrapping

As motivated in Section 2.3.3, a new LR phase unwrapping algorithm was devised to reduce the probability of phase discontinuities, called phase *glitches* [Coe01], arising in the unwrapped LR phase of a fine tuning subscan.

An example of a phase glitch induced by chopper cross-over is shown in Figure 2.8. Chopper cross-over was normally contained within the redundant DAQ cycles allocated for this purpose, as in Figure 2.4. Occasionally, a time lag in the synchronization of the choppers caused the cross-over between lasers to spill into the adjacent DAQ cycles allocated for Laser 1 or Laser 2, thereby corrupting the interferometer signals. This is shown for the SR in Figure 2.8a, where both or neither of the lasers illuminated the SR in the affected DAQ cycles. The corrupted LR signal created errors during LR phase extraction; either the wrong phase was extracted or an error flag was raised and the offending data point was removed. The wrong phase and/or the missing data

[12]The software was written in collaboration with Paul A. Coe.

[13]The original software saved the data by individually writing each array element to a file, rather than streaming the data to a file.

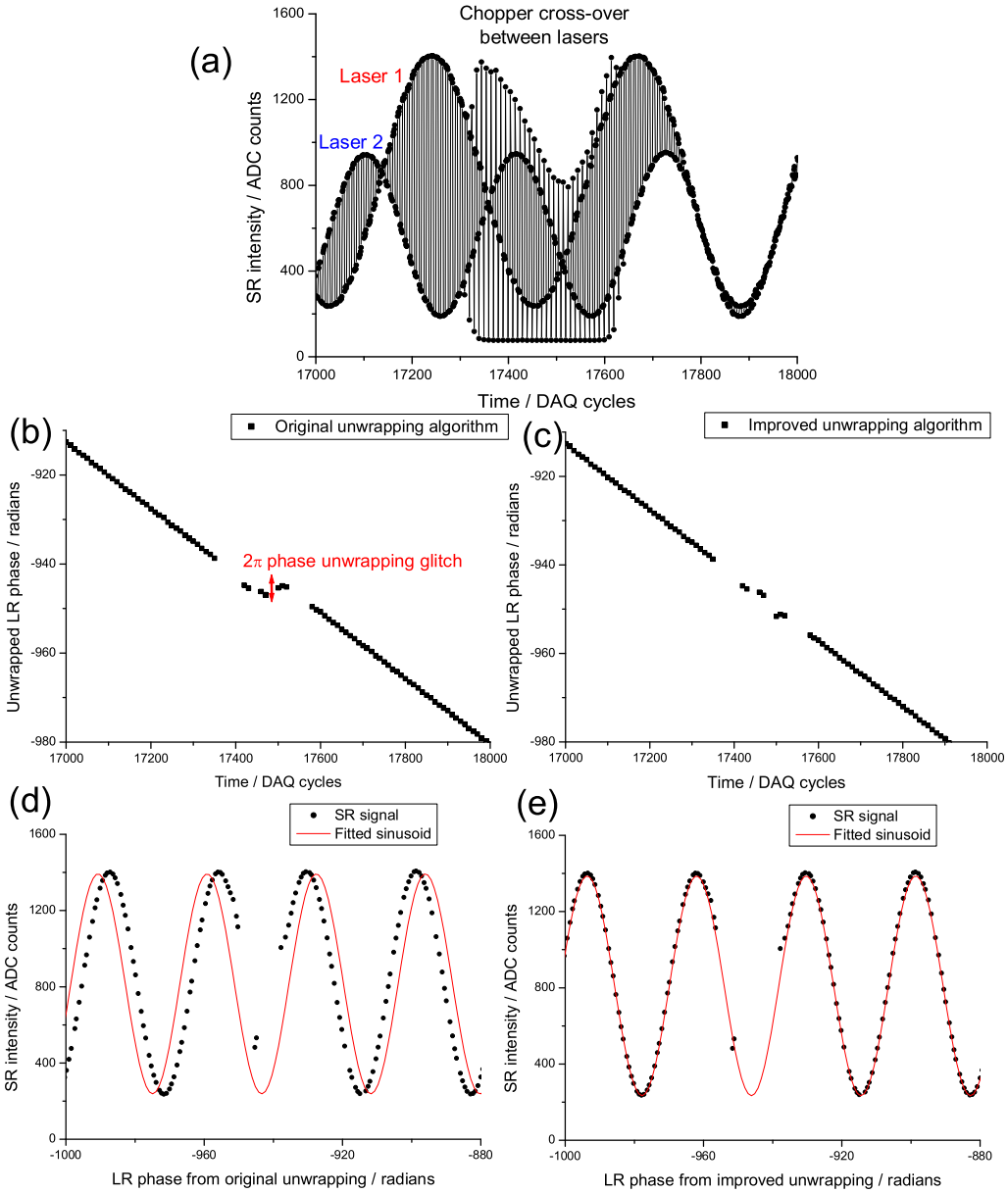

Figure 2.8: *An example of chopper cross-over (a) creating a glitch of 2π in the LR phase (b), when unwrapped with the original algorithm. When the SR intensity is plotted versus the LR phase (d), the LR phase jump creates a mismatch between the sinusoidal data on either side of the glitch, causing a misfit that renders the subscan useless. An improved phase unwrapping algorithm withstands the chopper cross-over (c), so that the fitted sinusoid (e) may be successfully used in linking.*

often foiled the original phase unwrapping algorithm [Coe01] and caused one or more phase glitches, such as the one in Figure 2.8b. The sinusoidal fits to all interferometers were severely degraded by such glitches, which in the absence of manual correction, rendered the subscan useless.

A new phase unwrapping algorithm was devised that was more tolerant to phase extraction errors, particularly to large groups of removed data points caused by temporary bursts of chopper cross-over or multi-moded lasing. In the majority of cases, such as that shown in Figure 2.8c, the new algorithm successfully unwrapped the LR phase that had contained glitches when unwrapped with the original algorithm. In these cases, the sinusoidal fits to the interferometer signals were found to be of sufficient quality for reliable linking, despite the missing or corrupted data.

The original and new unwrapping algorithms were both defeated by particularly awkward combinations of consecutive phase errors, induced by severe multi-moded lasing and/or chopper cross-over. In these exceptional cases, the subscan containing the phase glitch was inspected and the longest region of continuous phase was manually selected for the sine fitting analysis, so that linking could proceed as normal.

The ATLAS FSI system should have lasers less prone to multi-moded lasing than the lasers of the prototype FSI system. In future, the chopper synchronization should be improved or the redundant time allocated to contain chopper cross-over could be extended.

2.4.3 Automated Etalon Peak Counting and Verification

When linking the LR phase, the extrapolation between subscans relies upon the best estimate of the phase ratio of the etalon, q_{ET}, and on the number of etalon peaks, N, between a peak in each subscan, as in Equation 2.31. In previous work [Coe01], q_{ET} was estimated with sufficient precision to allow LR phase linking over ~ 3.5 THz, if the correct number of etalon peaks, N, were provided by hand. Failure to provide the correct number of etalon peaks caused all GLI lengths to be miscalculated. The number of etalon peaks was verified and if necessary, manually corrected, by repeating the linking process for different values of N and manually checking the linked length of the SR for compatibility with its previously estimated temperature dependent value [Coe01]. This was a tedious process for large data sets and typically there were ambiguous choices of N, for which the linked and estimated SR lengths were compatible. As a last resort, these ambiguities were resolved by counting the number of etalon peaks recorded during the coarse tuning between subscans. The narrow etalon peaks were sometimes missed during coarse tuning, due to the relative coarse tuning and sampling rates, leading to miscount errors.

New FSI software was written so that the choice of ADC channels recorded in the DAQ cycle could be different for fine and coarse tuning subscans. The etalon ADC channel was set to acquire data in almost every sequence of a DAQ cycle during coarse tuning, as in Table 2.1. The etalon peaks were counted by an algorithm that identified potential peaks when a threshold boundary was crossed by the etalon intensity data, but only incremented the peak count if sufficient tuning time had elapsed between adjacent potential peaks. The number of peaks, N_c, counted by this algorithm was used in the following automated procedure that reliably determined the correct number of etalon peaks, N.

Etalon peak count verification with the SR

The number of etalon peaks, N, was verified by repeatedly linking the SR between a pair of subscans for different values of N. Data for Laser 2 was linked first, as the linked result was needed to implement inter-subscan drift correction. The number of etalon peaks between subscans should have been zero for Laser 2 because it was tuned over the same frequency range in both subscans. To check the tuning range had not slipped by an etalon peak, the linking was repeated for N$= -1, 0, +1$. The value of N that minimized:

$$\left| q_{\mathrm{ET}}^{\mathrm{LINK}} - q_{\mathrm{ET}}^{\mathrm{XPD}} \right| \tag{2.54}$$

was taken, where $q_{\mathrm{ET}}^{\mathrm{XPD}}$ is the expected value of the phase ratio of the etalon and $q_{\mathrm{ET}}^{\mathrm{LINK}}$ was the value of the etalon phase ratio determined from the link.

The SR was then repeatedly linked for Laser 1 data, while the number of etalon peaks, N, was varied about the value estimated from the peak count[14] so that N=$N_{c\,-9}^{+4}$. The value of N that minimized the quadrature sum of the relative errors in the phase ratios q_{SR} and q_{ET}:

$$\left| \frac{q_{\mathrm{SR}}^{\mathrm{LINK}} - q_{\mathrm{SR}}^{\mathrm{XPD}}}{q_{\mathrm{SR}}^{\mathrm{XPD}}} \right|^2 + \left| \frac{q_{\mathrm{ET}}^{\mathrm{LINK}} - q_{\mathrm{ET}}^{\mathrm{XPD}}}{q_{\mathrm{ET}}^{\mathrm{XPD}}} \right|^2 \tag{2.55}$$

(where $q_{\mathrm{SR}}^{\mathrm{XPD}}$ is the expected SR length accounting for temperature)

was taken as the correct number of etalon peaks between the subscans and was used in subsequent linking of all GLIs.

[14] N_c also includes the sum of etalon peaks recorded in the relevant portion of each fine subscan between the two linked peaks. The search range was asymmetric due to a tendency to undercount the etalon peaks. Note N_c was always negative because Laser 1 tuned down in frequency.

2.4.4 Multiple Subscan Linking

As outlined in Section 2.4.1, the technique of multiple subscan linking was developed to improve the reliability and precision of FSI measurements. A set of fine tuning subscans were interspersed with periods of coarse tuning to extend the effective frequency range of the FSI scan, as shown in Figure 2.9. The different combinations of pairs of fine tuning subscans were linked sequentially, in order of increasing frequency interval separating the pair, as in Table 2.3. The final and largest link provided the GLI length.

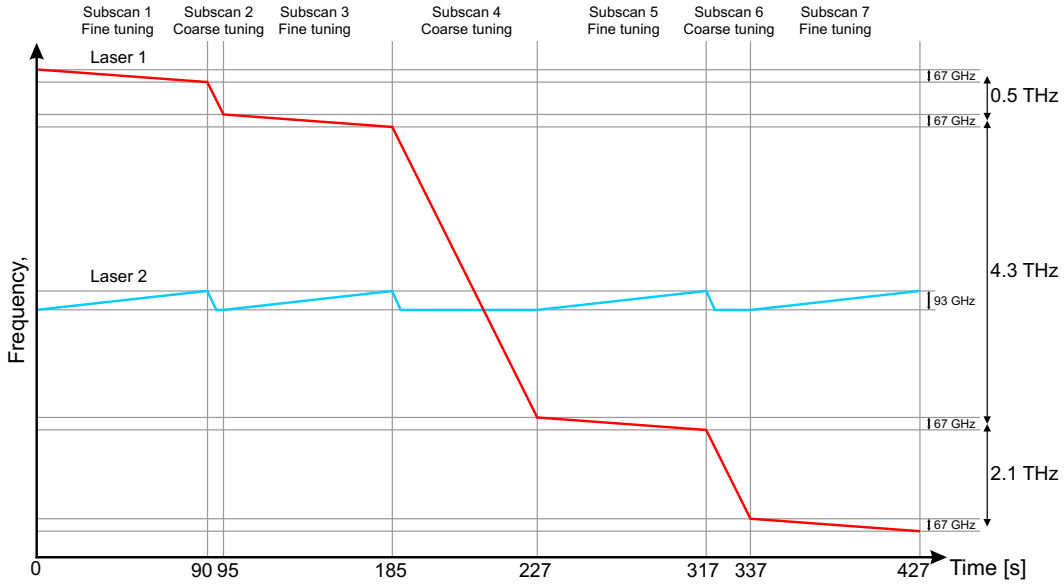

Figure 2.9: *A typical frequency tuning sequence for multiple subscan linking. The fine tuning ranges of both lasers are exaggerated.*

Link Order	Linked Subscan Pair	Set Linking Bridge / THz	Mean Etalon Peak Count
1	$1 \Longleftrightarrow 3$	0.5	-59.59
2	$5 \Longleftrightarrow 7$	2.1	-202.87
3	$3 \Longleftrightarrow 5$	4.3	-409.00
4	$1 \Longleftrightarrow 5$	4.8	-468.59
5	$3 \Longleftrightarrow 7$	6.4	-611.87
6	$1 \Longleftrightarrow 7$	6.9	-671.46

Table 2.3: *The 6 link permutations of subscan pairs from the 4 fine tuning subscans of Figure 2.9. The subscan pairs were sequentially linked in order of increasing frequency interval separating the pair (the linking bridge). The mean sum of etalon peaks counted between each subscan pair is shown for 56 FSI scans recorded with this coarse tuning pattern.*

The laser tuning pattern for multiple subscan linking was implemented without modifying the original clock and control card. Instead, the ability to pause the coarse tuning of Laser 1 mid sweep was exploited. The size of each link was controlled by adjusting the time intervals between pauses in coarse tuning. An estimated 0.24 nm s^{-1} linear coarse tuning rate was assumed, although the actual coarse tuning rate varied slightly due to non-reproducible motion of the piezo-actuated mirror in the laser head. These variations were the source of the differences between the linking bridge sizes set and those measured by the 10 GHz etalon, as in Table 2.3.

When linking the GLI phase between a pair of subscans, the extrapolation relied upon the best estimate of the GLI phase ratio, q_0^{AV}, in Equation 2.39, and also on the link corrected LR phase difference between the subscan mid-points, $\Delta\Phi_{\mathrm{AB}}$. If the LR phase had been successfully linked, then errors on $\Delta\Phi_{\mathrm{AB}}$ were relatively small. The extrapolation errors were dominated by the error on q_0^{AV}, which was due to the residual phase error in the GLI intensity sine fit. If the extrapolated phase is within $\pm\pi$ of the locally measured phase in the

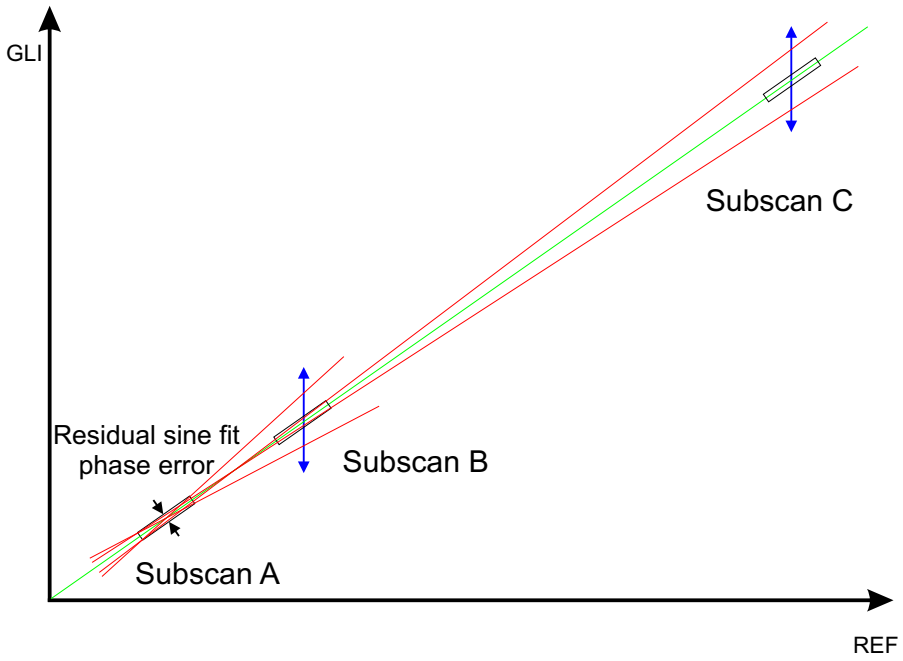

Figure 2.10: *An example of multiple linking improving the estimate of the phase ratio,* $q = \frac{\Delta\Phi_{\mathrm{REF}}}{\Delta\Theta_{\mathrm{GLI}}}$. *The extrapolation errors from subscan A to subscan B lie within* $\pm\pi$ *of the local phase of subscan B, so the subscans may be linked. The improved estimate of the gradient allows a much larger link between subscans A and C (and also between subscans B and C, although this is not shown).*

second subscan, as in Figure 2.10, then after linking the first two subscans, the phase ratio estimate drastically improves to allow a much larger subsequent extrapolation to another subscan.

The linking bridge size between adjacent subscans did not need to progressively increase. The subscan pair permutations were sorted prior to linking in order of increasing bridge size, by checking the sum of etalon peak counts between each subscan pair. The coarse tuning pattern was chosen to include at least one very short, safe link (<THz) well within $\pm\pi$ and to ensure the remaining progressively larger links were within the $\pm\pi$ extrapolation limit, with the minimum number of links to reduce the total measurement time and hence adverse effects of interferometer drift.

56 FSI scans were recorded over 9.5 hours using the tuning pattern shown in Figure 2.9 and were linked in the order shown in Table 2.3. The 6 estimates of the SR length from each link permutation of subscan pairs, are shown for all FSI scans in Figure 2.11a, plotted against the scan averaged temperature of one invar rod of the SR. The SR had a negative CTE due to its construction, as found in previous work [Coe01] in which the expansion was assumed linear.

Linear and second order polynomials were fitted to the SR length data for each link permutation. The second order polynomial fit to data from the largest link is shown in Figure 2.11b and the standard deviation of the fit residuals was 7 nm for the 170 mm SR length (0.04 ppm). This represents the statistical precision in determining the SR length when using that specific link permutation.

The standard deviations of the residuals from either fit to data from each link permutation are shown in Figure 2.11c. This length error, $\varepsilon_{\mathrm{GLI}}$, was found to reduce with linking bridge size. The reduction approximately follows an inverse law as expected from:

$$\frac{\varepsilon_{\mathrm{GLI}}}{\mathcal{D}_{\mathrm{GLI}}} \propto \frac{\varepsilon_{\Theta}}{\Delta\Theta_{\mathrm{LINK}}} \tag{2.56}$$

where ε_{Θ} is the mean residual phase error on the GLI intensity sine fit. Deviations from this form are thought to be due to fluctuations in the residual sine fit phase errors of different subscans.

Note that the mean of the lengths from each link permutation do not coincide, indicating there is also a systematic error for each link. Importantly, this systematic error does not reduce with increasing linking bridge size: even though it generally appears to do so in Figure 2.11d, by the narrowing in of the mean SR length from each link permutation data set. A remarkable exception is that the link size increases from link 3 to link 4 but the systematic error does not reduce. The systematic error was found to be correlated with the difference in mean frequencies of lasers for each link permutation, as shown

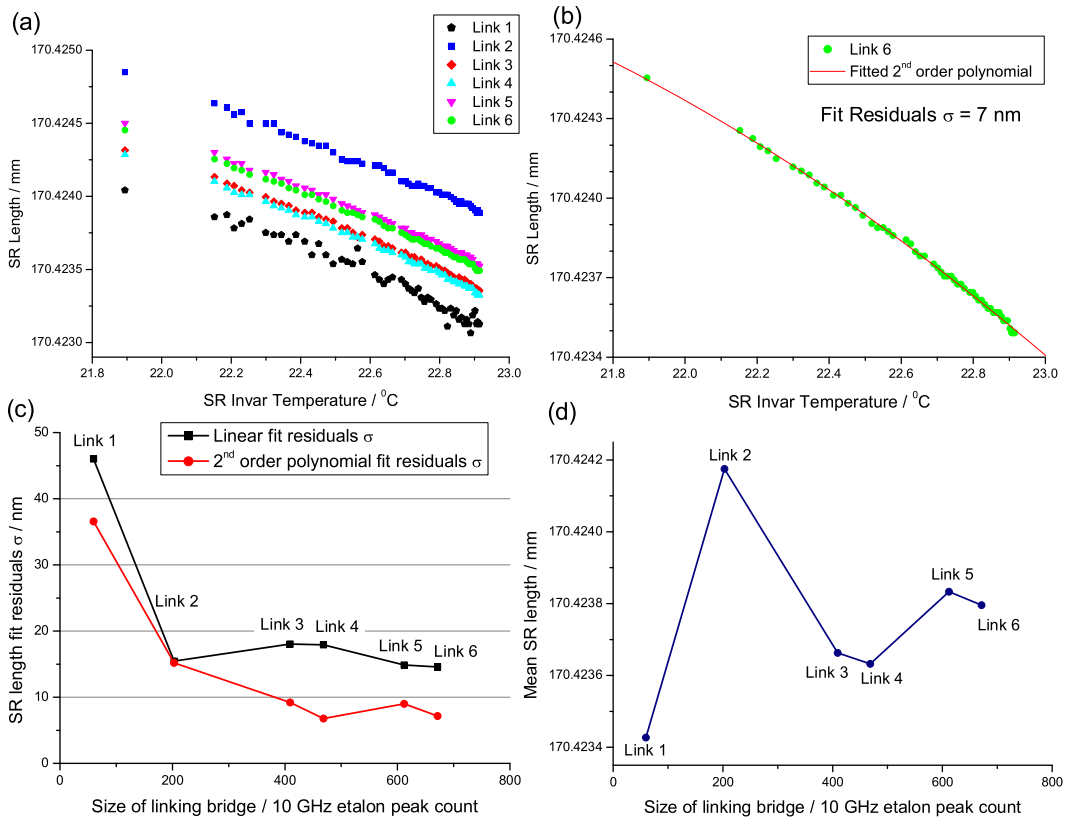

Figure 2.11: *An example of multiple subscan linking for the short reference interferometer. The first scan was recorded 1.6 hours before the second.*

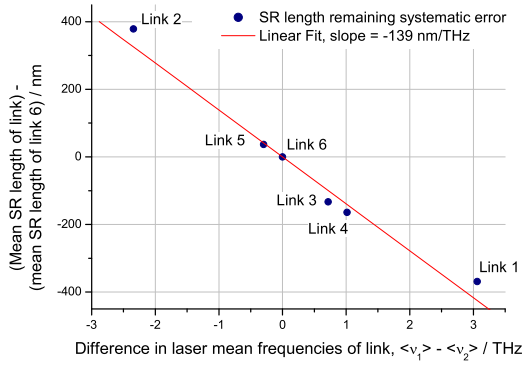

Figure 2.12: *The systematic error remaining on the SR linked length in Figure 2.11d is correlated with the difference in mean laser frequencies of a given link. This is thought to be due to a residual interferometer drift error.*

in Figure 2.12. This correlation indicates that the remaining systematic error could be due to an interferometer drift error that has not been completely cancelled. It is thought an error in the absolute frequency ratio of the two lasers, $\frac{\nu_1}{\nu_2}$, which is included in the drift correction of Equation 2.50, could be the source of this error.[15] Crucially, the systematic error reduces with the difference in the mean laser frequencies, so by setting these to be equal for the largest link, the remaining systematic error can be significantly reduced and almost eliminated.

The ATLAS FSI system only requires repeatable, precise measurements. The linking pattern used during the calibration of the FSI system with the X-ray survey (see section 5.5) and during operation in ATLAS should ideally be kept the same, to reduce any systematic differences incurred by a change of linking pattern. Also, the mean laser frequencies for the largest link should be equalized to reduce systematics. It would be preferable to repeat the calibration with a variety of linking patterns, in case the operational linking pattern needed to be changed at a later date.

2.4.5 Multiple, Simultaneous GLI Measurements

In previous work [Coe01, Mit02] only single GLI length measurements were performed. The ATLAS FSI grid contains 842 GLIs that should be simultaneously measured. The hardware of the prototype FSI system was upgraded to include remote multiple GLIs, configured as a geodetic grid, that were all simultaneously powered and read out, as discussed in Chapter 3. New software was written to acquire and analyse data from the multiple GLIs, as required for the grid measurements in Chapter 4.

2.4.6 GLI Measurement Precision

The measurement precisions for GLIs were typically ∼200 nm as calculated in Chapter 4. The precision was much worse than for the SR, primarily due to the low power return signals in the GLI design and unwanted W-mode reflections that induce residual phase errors when sine fitting. The precisions were well within the 1 μm precision required for the ATLAS FSI system.

[15] Since this was originally written, further data sets have been investigated showing a near identical correlation under various drift conditions. A path difference in beam-splitter(s) of the SR and/or LR of ∼mm would generate the observed -140 nm/THz shift in linked SR length. The frequency dependence of the beam-splitter glass refractivity causes an effective change in interferometer OPD with frequency. [Coe04]

2.4.7 Recommendations for Future Improvements

In previous work [Mit02] the laser fine tuning rate was linearized and in most cases studied an improvement in the measurement precision was observed. Laser linearization was not implemented for multiple subscan linking because the same voltage ramps were applied in each subscan to tune the lasers but the tuning curves were either different in each subscan or non-reproducible. The system could be modified to apply a different non-linear voltage ramp to the different fine tuning subscans, to linearize the tuning rate.

In a few FSI scans, a poor GLI intensity sine fit in one subscan caused the linking sequence to fail because of a poor estimate of the phase ratio, q, even though the other subscans could otherwise be linked with the previous best estimate of q. A method of identifying the quality of the phase ratio estimate for a given link should be developed, so that links which worsen the estimate could be omitted from the linking sequence. The tuning pattern could be modified to allow redundant links for this purpose.

The sine fitting method was generally robust, although very occasionally ($< 1\%$) the fit failed for a given seed of the interferometer length, even though a good fit was found for seed lengths very close to the original. A more robust sine fitting method should be sought or the sine fit quality should be automatically checked and the GLI intensity refitted with a slightly different seed if the fit failed (within a limited number of iterations).

The unwanted W-mode reflections in the quills were normally the largest source of error in the residual phase errors on the sine fit and limited the measurement precision for a given linking bridge. Software methods for overcoming the effects of unwanted reflections in sine fitting should be investigated.

Linking between lasers

A possible future improvement might be to link between lasers using a tuning sequence similar to that shown in Figure 2.13. The final and largest link would be between the two lasers in subscan 1. Such a scheme may create new systematic errors, perhaps due to the opposite laser tuning directions and piezo balancing errors. However, the scheme would have the triple advantage of reducing the total measurement time, allowing more link permutations and eliminating inter-subscan drift and associated errors from the final link. Further investigation is required.

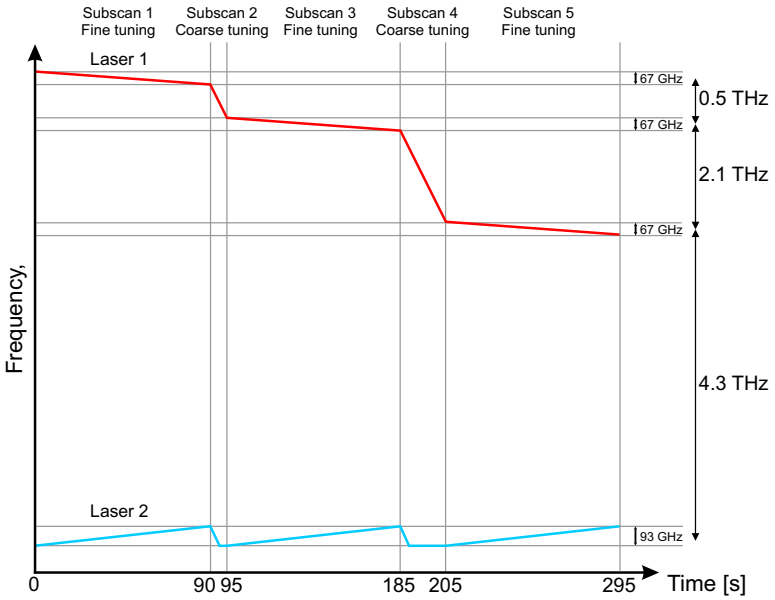

Figure 2.13: *A possible tuning sequence for multiple subscan linking. The same total link size as in Figure 2.9 could be achieved in less time by linking between lasers in the first subscan. This would eliminate inter-subscan drift errors in the final linked length.*

2.5 Spurious Reflections

2.5.1 Introduction

Spurious reflections within the ATLAS SCT that couple light into the return fibre of a GLI cause beating in the interferometer signal [Coe01]. During sine fitting of the GLI intensity, the unwanted interference often degrades the sine fit and causes large residual phase errors, which are the dominant error in the final linked length. Disentangling the correct GLI length becomes especially problematic, when the optical path length traversed by the spuriously reflected light is similar to that of the intended GLI. Two types of potential spurious reflections in the ATLAS FSI system were investigated as summarized in Sections 2.5.2 and 2.5.3.

2.5.2 Glancing Incidence Reflections

In the ATLAS FSI grid many GLIs run close and often parallel to the interior surface of the carbon-fibre cylinders that constitute the ATLAS SCT Barrel. Reflections are particularly strong for materials at glancing angles of incidence. The longest GLIs that span the 1.5 m barrel length are particularly susceptible

to glancing incidence reflections. These potential reflections were replicated[16] by setting up a 1.5 m mirror plane adjacent to a GLI, with the same proximity as in the ATLAS design. The quill at one end of the mirror plane targeted a retro-reflector mounted in a prototype jewel, as shown in Figure 2.14.

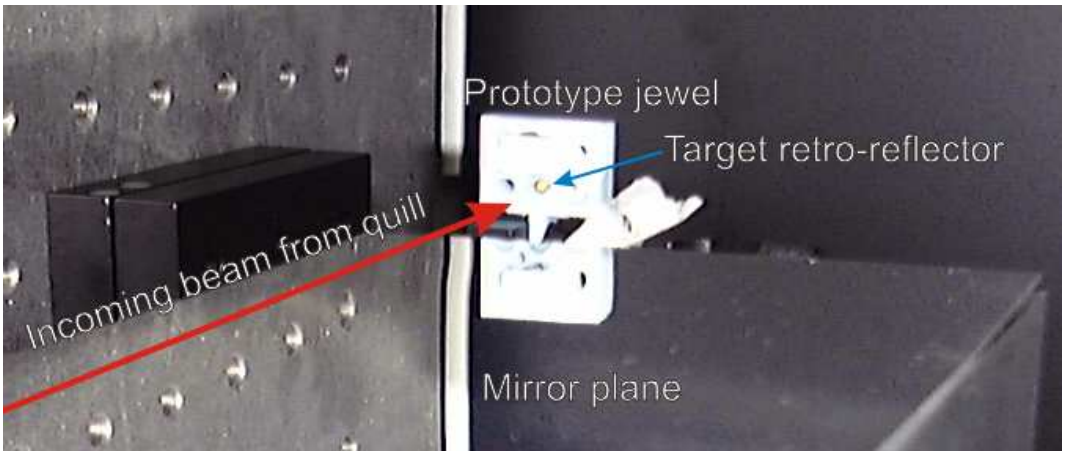

Figure 2.14: *A prototype FSI jewel that held the retro-reflector targeted by a quill at the far end of the mirror plane. The proximity of the mirror plane matches that of the carbon fibre surface in the ATLAS SCT.*

A set of consecutive FSI scans were recorded with and without the mirror plane in place. The GLI length determined from each FSI scan is plotted in Figure 2.15 against the scan averaged temperature of the optical table to which the GLI was mounted. The expected thermal expansion of the steel, is shown.

The presence or absence of the mirror surface made no difference to the measured lengths. The standard deviation of the residuals differences between the measurements and the thermal expansion trend is 990 nm. Further details can be found elsewhere [Coe02b].

2.5.3 GLI Protection Tubes

At the ends of the ATLAS SCT Barrel, FSI measurements from one cylinder to another are protected by tubes, to prevent services obscuring the light path (as in Figure 5.14). Due to the space constraints, these tubes are narrow compared to the laser beam propagating inside them. The tubes had to be carefully designed to allow light from the quill directly to and from the retro-reflector, while suppressing any spurious reflections via the tube walls that might couple light to the return fibre.

[16]The glancing incidence reflections test was performed in collaboration with Paul A. Coe.

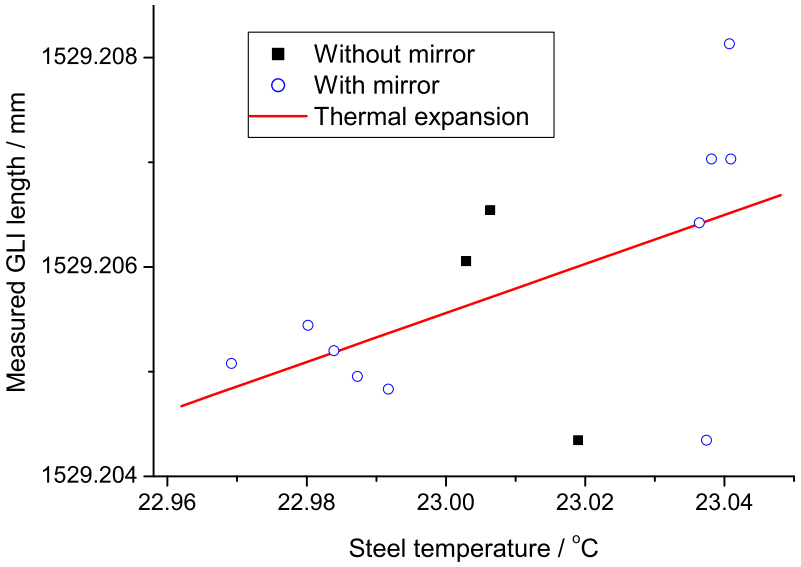

Figure 2.15: *A 1.5 m GLI measured in the presence and absence of glancing incidence reflections from a plane mirror adjacent to the GLI. The expected thermal expansion trend is shown for comparison. The standard deviation of residuals of the measurements from this trend is 990 nm.*

Tube design evolution

Several designs were tested as summarized below and detailed in [Gib01a]. The main steps in the design evolution are shown in Figure 2.16. Each tube was tested by placing it between a quill and retro-reflector that formed a GLI and comparing the interferometer fringes with those when no tube was present. A CCD camera was also used in place of the retro-reflector, to examine the interference pattern produced in the far field by reflections from each tube.

Two problems were encountered with various tube designs:

1. For tubes with a smooth flat side, strong reflections from the side wall led to extra path lengths in the interferometer that were similar in length to the direct GLI path. These produced slow beating with the direct path signal in the interferometer fringes, dramatically degrading the quality of fit. The far field image showed Lloyd's Mirror type interference fringes from the side walls, as in Figure 2.17, confirming the existence of strong side reflections.

2. Tubes with a ribbed-baffle design eliminated the side reflections, but introduced direct back scattering from the ribbing. These unwanted reflections also degraded the quality of fringes for the desired GLI signal.

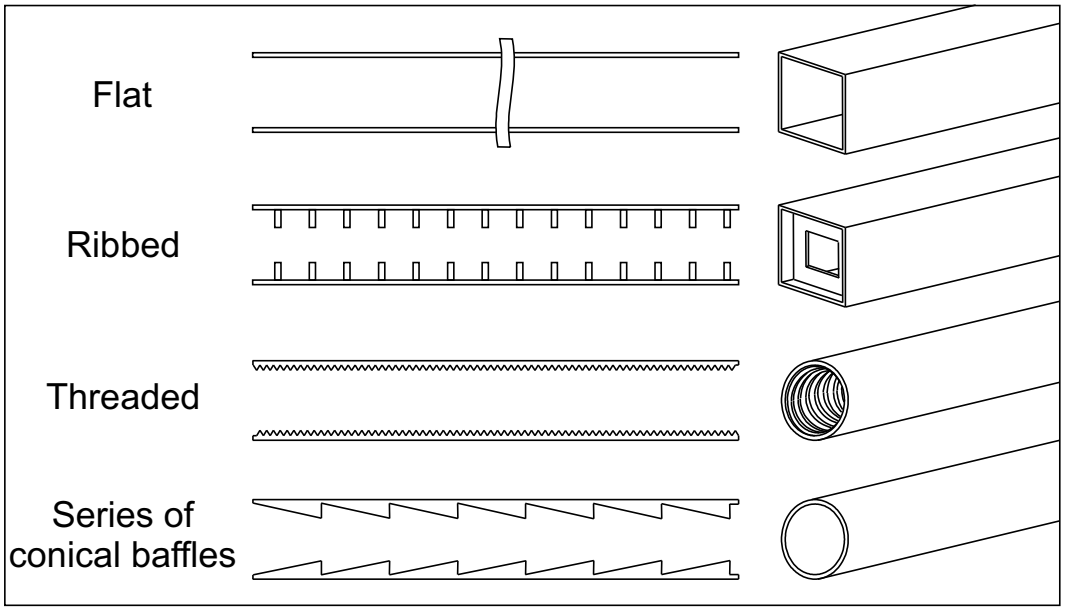

Figure 2.16: *Light path protection tube design evolution.*

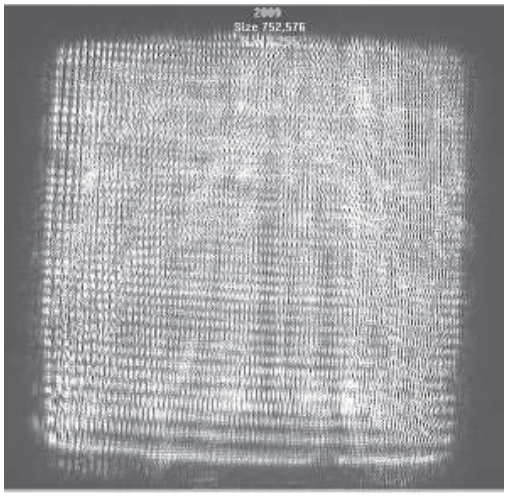

Figure 2.17: *CCD image of the far field pattern produced via reflections from a flat sided tube of square cross-section.*

Some tubes with internal threading were tested, with the aim of avoiding side reflections and back scattering. The CCD image still showed modest side reflection, thought to be from the curved apex of the threads, and some beating was also present in the interferometer fringes, corresponding to back scattering from threads near to the delivery fibre.

Tube design convergence

Threads were found to be ineffective baffles, which did not reduce the side reflections enough. A better design can be made by altering the gradient of the baffles to be steep enough to prevent side reflections but shallow enough to avoid back scattering.

A tube comprising a series of conical baffles, which were shallower than the threads or ribbing, prevents light being directly back scattered. Side reflections are also suppressed, due to the angle of reflection increasing as the light is reflected down the tube, as shown in Figure 2.18.

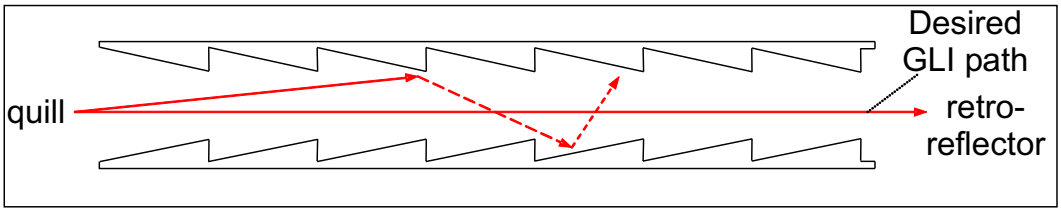

Figure 2.18: *Suppression of back scattering and side reflection using a series of conical baffles.*

A few iterations were required to achieve an appropriate gradient of the baffles. A 65 mm tube made from a sequence of conical baffles was tested in a GLI. A consecutive sets of FSI scans were recorded with and without the tube and the measured length of the GLI is shown in Figure 2.19. The presence of

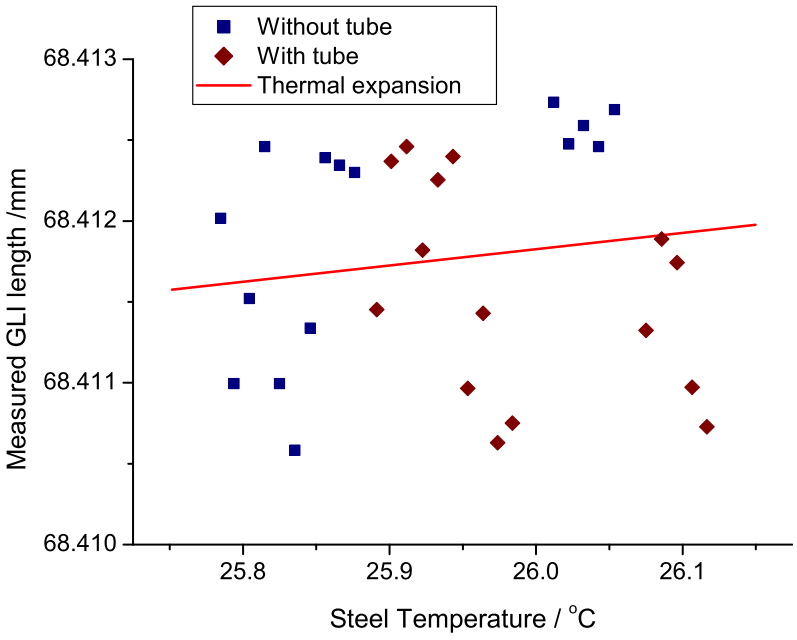

Figure 2.19: *The measured GLI length with and without the protection tube.*

the tube was found to induce slight beating in the interferometer signals, but made very little difference to the measured length. The standard deviation of the measured lengths is 0.7 μm.

The design of the GLI protection tubes for ATLAS is shown in Figure 2.20. The above test should be repeated for tubes of the final material (PEEK) when available. A similar conical baffle design should be implemented in the T-posts, shown in Figure 5.14.

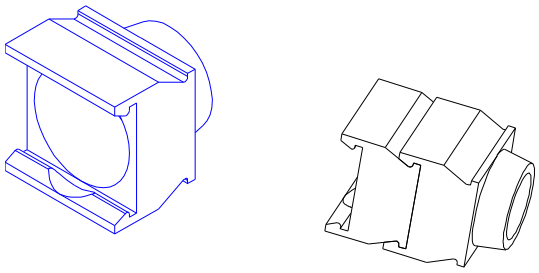

Figure 2.20: *The design of the GLI protection tubes consists of repeating units that are connected to form a series of conical baffles to suppress internal reflections.*

2.6 Rapid Optical Path Changes

2.6.1 Turbulence

Rapid variations in the refractive index of the gas within the ATLAS SCT are a potential source of error in FSI measurements. The tolerance of an FSI measurement to the effects of rapid refractive index variations was investigated with some very basic tests.

The tip of a soldering iron was positioned just beneath the line-of-sight of a GLI, and positioned midway between the quill and the retro-reflector. The axis of the soldering iron was initially perpendicular to the line of sight, so that only about a tenth of gas volume within the 252 mm GLI was drastically disturbed by thermally driven convention currents and turbulent gas flows.

Two groups of 5 FSI measurements were made before the heat source was applied. The soldering iron was then set to 330 °C for five further FSI measurements, increased to the maximum 420 °C for another 5 measurements, and then the soldering iron was switched off but was not removed and 5 further FSI measurements were recorded as it cooled.

Throughout the measurements the temperature of the gas near to but not directly above the soldering iron was monitored, giving the readings shown in Figure 2.21. The variations within the GLI itself were much greater.

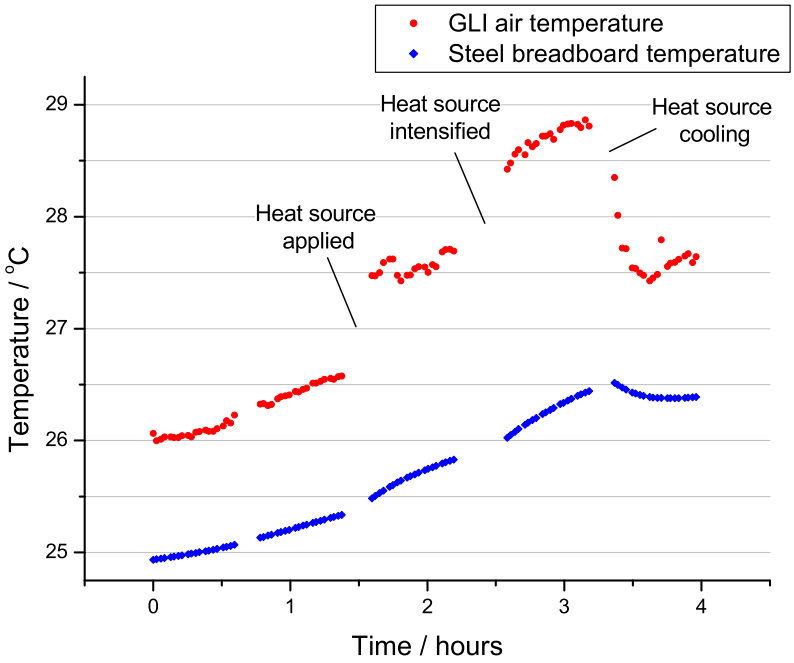

Figure 2.21: *The variation in gas temperature within a GLI, as an intense heat source was applied from beneath the line-of-sight. Each point represents the average temperature during one FSI fine tuning subscan, of which there were four per FSI scan. The gas thermometer was near to but not above the heat source.*

The interferometer signal was severely degraded when the heat source was applied, as shown in Figure 2.22. The rapid refractive index variations lead to optical path length changes that are clearly visible as the local change in wavelength of the FSI fringes.

Despite this horrendous, degraded signal the FSI length measurement precision is barely affected, as shown in Figure 2.23. The χ^2 profile of the sine fit is generally degraded but does exhibit the deep, distinct minimum required for precise FSI measurements. Only the fitted wavelength and phase affect the linked result. It is thought that because the variations are random and on short timescales compared to the FSI subscan, any pull of the phase during the sine fit is likely to average out.

A second test was performed with the axis of the soldering iron parallel to and directly beneath the axis of the GLI, so that the significant majority of the gas within the GLI was thermally agitated. In this very severe test, the χ^2 profile of the fitted signals often lost its distinct minimum and in most cases the FSI measurements failed to link. More advanced software methods

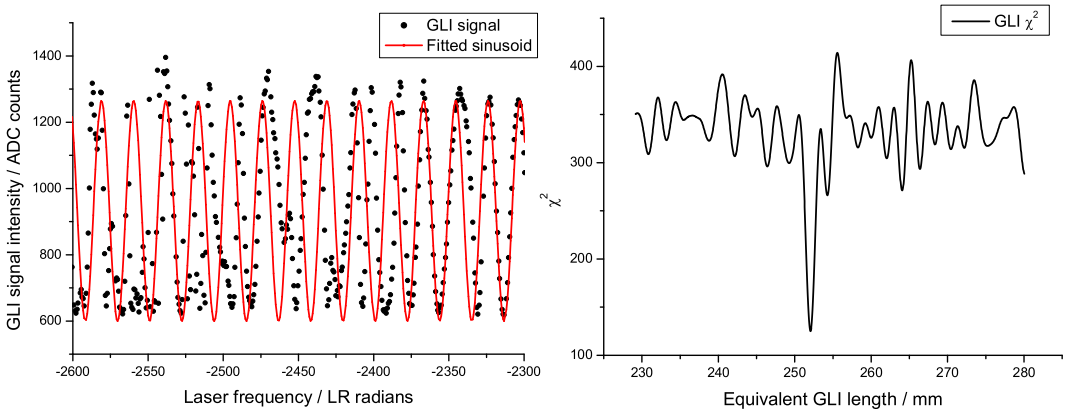

Figure 2.22: *GLI signal degradation induced by an intense heat source under the GLI creating turbulent gas flow and convection currents. The rapid refractive index variations change the optical path length on short time-scales. However, the χ^2 profile of the fitted sinusoid exhibits a deep, distinct minimum at the correct GLI length.*

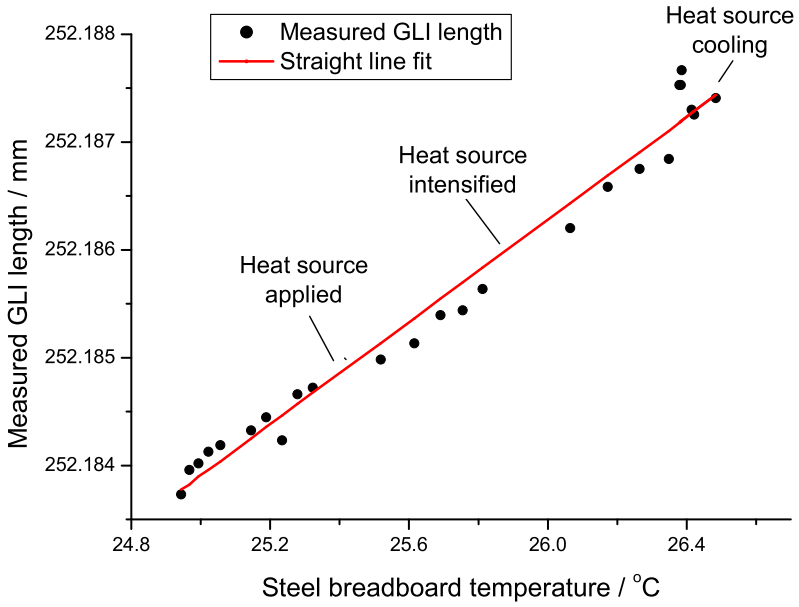

Figure 2.23: *The measured length of a GLI was barely affected by degradation of the GLI signal, as in Figure 2.22, induced by a 420°C heat source. The observed thermal expansion of 9.5 ppm/K is compatible with that for the steel breadboard. The residuals from the best fit have a scatter of 198 nm.*

to overcome this could be investigated.

Nevertheless, the refractive index variations in the ATLAS SCT are expected to be very much smaller than the tests here, in which a thermal gradient of ∼400 °C was applied. In ATLAS the difference between the module temperatures and the gas within the SCT should be only $\lesssim$25 °C [Due01]. Although the total gas volume within each ATLAS GLI will be exposed to refractive index variations, the FSI measurements are expected to be tolerant to these effects.

2.6.2 Vibrations

Vibration of the ATLAS SCT support structure is expected to induce rapid phase shifts in the interferometer signals. In this work no adverse effects from vibrations under normal working conditions were observed. The tolerance of FSI measurements to vibration is discussed in more detail elsewhere [Coe01, Coe02b].

2.7 Retro-reflector Tests

Retro-reflectors are an essential part of the FSI system. The prohibitive cost of suitably small, commercial retro-reflectors stimulated the in-house manufacture of the retro-reflectors [Huf01]. The variable production quality required suitable tests to be developed for quality assurance.

The WHIPM system [Huf01] had been developed for this purpose, which allowed the optical surface profile of retro-reflectors to be examined via an interferogram, an example of which is given in Figure 2.24.

The WHIPM system provided an excellent measure of the surface profile, which was found to be (see below) strongly correlated with the suitability of a retro-reflector for FSI measurements. However, some retro-reflectors that passed the WHIPM test were unsuitable for FSI (see Section 3.5.1).

A complementary test was developed, in which the far field pattern from the retro-reflector was examined. A retro-reflector was placed in a special holder that allowed it to be rotated to any orientation about its axis of symmetry and also rotated about the apex of the retro-reflector by 0° to 12.5° with respect to an incident laser beam from a quill via a beam-splitter, as shown in Figure 2.25. A CCD recorded the far field image in the plane equivalent to that of the quill. Importantly, the intensity at the centre of this image indicated the intensity at the quill return fibre.

An ideal retro reflector should have a bright central spot at all rotation angles permitted by the holder. The majority of tested retro-reflectors did not

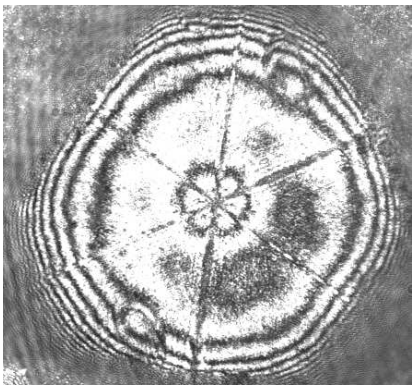

Figure 2.24: *The interferogram of a retro-reflector recorded with the WHIPM system. The number of fringes across the figure quantifies the flatness of the optical profile.*

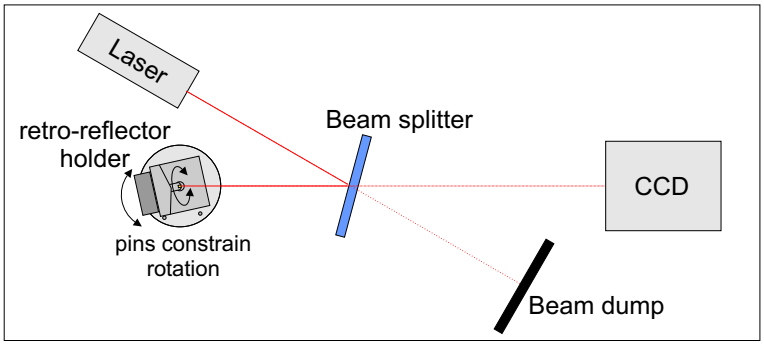

Figure 2.25: *The configuration for the far field retro-reflector test.*

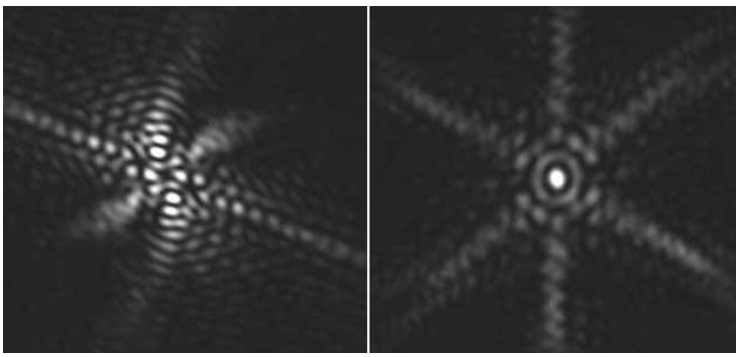

Figure 2.26: *The far field images from a poor and very good quality retro-reflector, distinguished by the absence or presence of a central bright spot, which indicates the intensity at the return fibre of a quill.*

meet this requirement. The quality was assessed primarily from the central spot intensity at the extreme orientations of the holder. The far field images from *poor* and *very good* quality retro-reflectors are shown in Figure 2.26.

The far field test results were strongly correlated with the WHIPM test results, except that some retro-reflectors that passed the WHIPM test were rejected by the more stringent far field test.

The signal in a functioning GLI was significantly diminished by reorienting a poor quality retro-reflector about its axis, when angled by $\sim 10^{\circ}$ to the beam from the quill. The far field test accurately predicted the orientation in which a poor quality retro-reflector resulted in the success or failure of FSI measurements. This test underlies the quality assurance for the retro-reflectors installed in ATLAS.

2.8 FSI Component Design

The design of components for the ATLAS FSI system takes into account the reflection issues discussed in Section 2.5. The quills and retro-reflectors are accurately positioned in small support structures called jewels. The jewel designs avoid spurious reflections by angling flat surfaces away from the incoming beams of light targeting that jewel. The angles were also chosen to avoid reflections that couple light from one quill to another quill. Where the jewels hold more than one retro-reflector, baffles were added to the design to prevent more than one retro-reflector being viewed by any quill, as in Figure 2.27.

The quills are pre-aligned in a jig as shown in Figure 2.28 and tested to check that no spurious reflections are present in the interferometer signals before the jewels are installed in ATLAS.

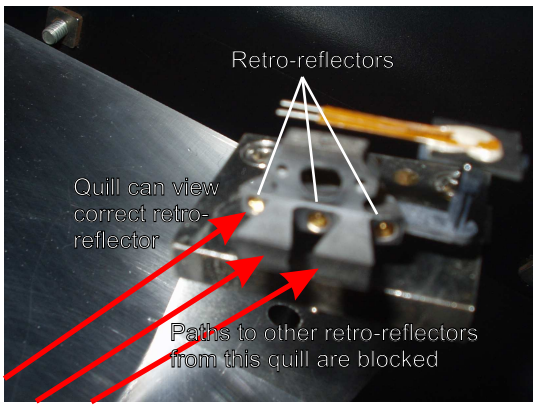

Figure 2.27: *An ATLAS FSI jewel that supports three retro-reflectors. Small baffles prevent more than one retro-reflector being viewed by any quill.*

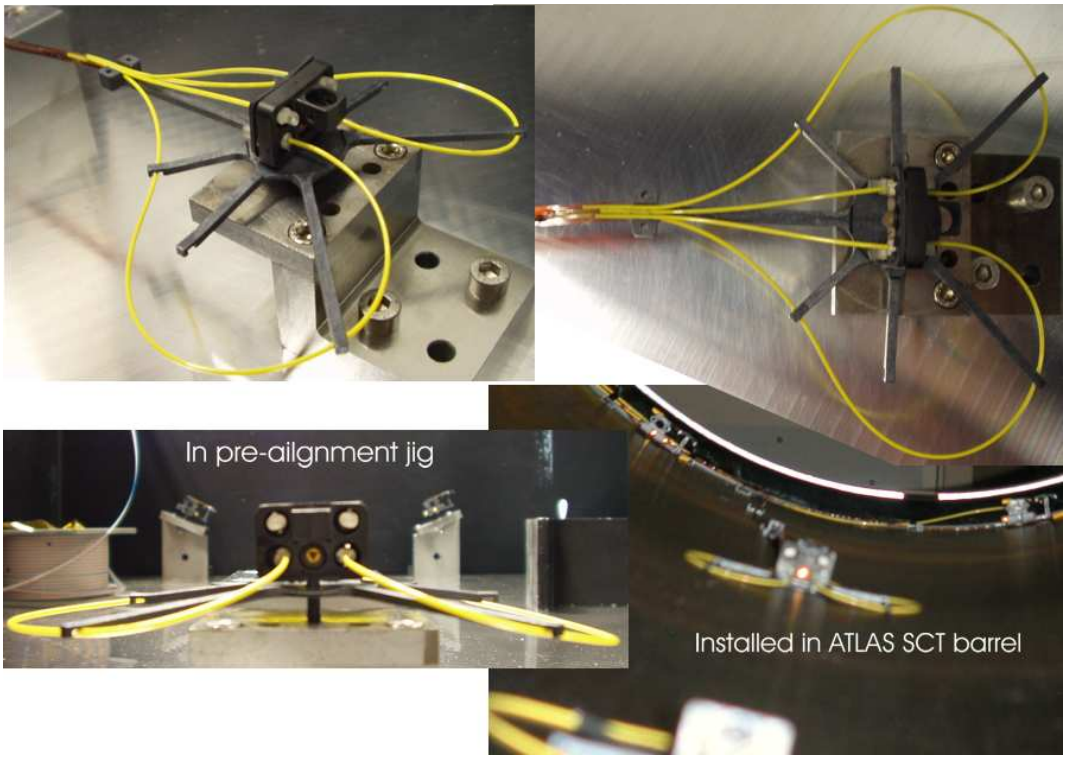

Figure 2.28: *An ATLAS FSI jewel, called a 'scorpion', is pre-aligned in a jig that replicates the layout of the ATLAS SCT barrel.*

2.9 Conclusion

The technique of Frequency Scanning Interferometry has been developed to make robust, precise length measurements in the hostile, inaccessible environment of the ATLAS SCT. This work contributes important advances in the FSI technique, which improve the reliability, precision and scope of the measurements.

The key technique of *multiple subscan linking* has been implemented in the prototype FSI system. With this technique, the length of a 170 mm reference interferometer has been measured to a precision of 7 nm (0.04 ppm). The precision has been shown to improve as the frequency interval of the total scan is increased. Some residual systematic errors, which are thought to be due to interferometer drift, were greatly reduced and almost eliminated by ensuring the mean laser frequencies are set to be equal.

The prototype FSI system hardware and software was upgraded to measure multiple GLIs for the first time, as detailed in the next chapter.

The FSI measurements have been shown to be tolerant to several types of spurious effects that might degrade the interferometer signals in ATLAS. Methods to suppress unwanted spurious reflections were developed and incorporated in the design ATLAS FSI components. In future, software methods to combat the effects of signal degradation on the measurement precision should be investigated.

Chapter 3

Prototype FSI Grid Set-up

3.1 Introduction

The previous chapter focused on FSI applied to individual distance measurements. However, multiple, simultaneous measurements are required for the ATLAS SCT alignment system. This chapter presents the development of the first, prototype FSI geodetic grids that employ multiple, simultaneous distance measurements between grid nodes.

Geodetic Grids in ATLAS

The purpose of the FSI alignment system is to measure shape deformations of the ATLAS SCT. This will be achieved by arranging the measurements as a **geodetic grid**[1]: a network composed of a number of nodes and measurements between nodes, each node generally being common to several measurements. The grid shape is reconstructed by combining the measurements to solve for the grid node co-ordinates. How the grid has been configured to fulfill its role in ATLAS, is addressed with detailed simulations in Chapter 5. The grid reconstruction software used for these simulations was validated with data from the prototype grids described below.

Demonstration System

Several simple prototype grids were constructed to demonstrate the high precision shape reconstruction possible using FSI. A prototype grid with square geometry was built to investigate node co-ordinate measurement in two dimensions. Subsequent adaptation to a tetrahedral geometry enabled three dimensional co-ordinate measurement to be examined.

[1] also called a geodetic network

A prototype grid formed part of the demonstration system, described in Section 3.2, that allowed the grid to be powered and read out in a controlled environment. Section 3.3 describes the construction of the prototype grids and how they were configured to improve the study prospects. The preliminary alignment of the grid components is described in Section 3.4. Initial trials and optimization of FSI signals from the grid are discussed in Section 3.5, with an emphasis on issues related to the measurement of the ATLAS FSI grids.

This chapter concerns only the set-up of the prototype grids. The investigation of prototype FSI grid shape measurement is the concern of Chapter 4.

3.2 Demonstration System Set-up

Each prototype grid was evaluated in a mock-up of the ATLAS FSI system. This was formed by connecting the dual-laser and reference interferometer system [Coe01], previously described in section 2.3.2, to the grid testing apparatus shown in Figure 3.1. An optical fibre conveyed light from the former to the latter, to power the prototype grid. A fibre splitter-tree allowed all grid line interferometers (GLIs) to be illuminated simultaneously. The GLI signals were read out by avalanche photodiodes (APDs).

The grid testing apparatus was arranged so that the fibre splitter-tree and APDs were remote from the FSI grid, as in the ATLAS FSI system. The components were housed in two compartments representing the ATLAS equipment (USA15) and detector (UX15) caverns. The grid was measured remotely via the delivery and return fibres, shown with yellow and orange furcation tubes respectively in Figure 3.1.

3.2.1 Environmental Chamber

The grid components were mounted on a Newport steel optical breadboard. The breadboard was insulated with 50 mm high-density polystyrene foam to reduce thermally driven motion of the grid nodes. During FSI measurements, the insulation was completed with a foam lid to reduce turbulence and slow down temperature changes of the air within the GLIs. A small hinged portion of the lid allowed access to the grid with minimal disturbance to the air.

The temperature of the optical breadboard and of the air between the grid nodes was monitored with type PTR 100 platinum resistance thermometers, read out with a custom built low noise circuit [Coe01]. The grid air temperature measurement allowed refractive index corrections to the optical path length of the GLIs by the method described in Section 2.3.3. The air pressures within

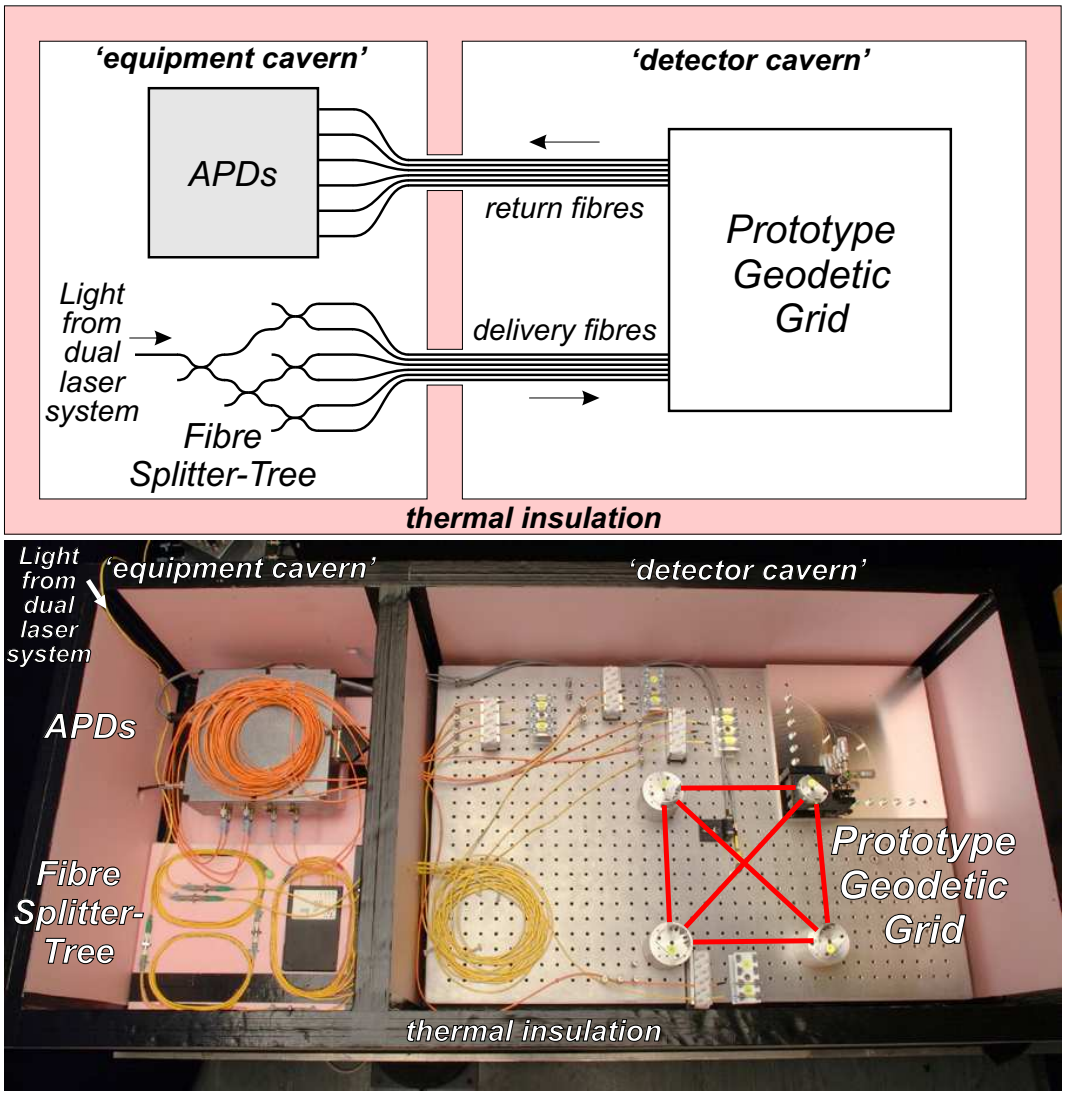

Figure 3.1: *Overview of the grid testing apparatus. The arrangement was a mock-up of the ATLAS system, in which the grid was powered and read out with APDs remotely via optical fibres.*

the reference interferometer system and the GLIs were assumed equal because they shared the same laboratory air.

3.2.2 Fibre Splitter-Tree

Several commercially available optical fibre splitters were connected to form the fibre splitter-tree shown in Figure 3.2. This divided the fibre coupled light from

the dual-laser system between the multiple GLIs. The optical power conveyed by and lost within each branch of the tree was estimated by systematically measuring the light emerging from each connector and quill. These measurements, presented in Figure 3.2, were taken while the grid was illuminated with Laser 1 at typical settings. The power reaching each GLI from Laser 2 was typically one third that of Laser 1. The fibre-splitter insertion losses were consistent with those specified by the manufacturer.

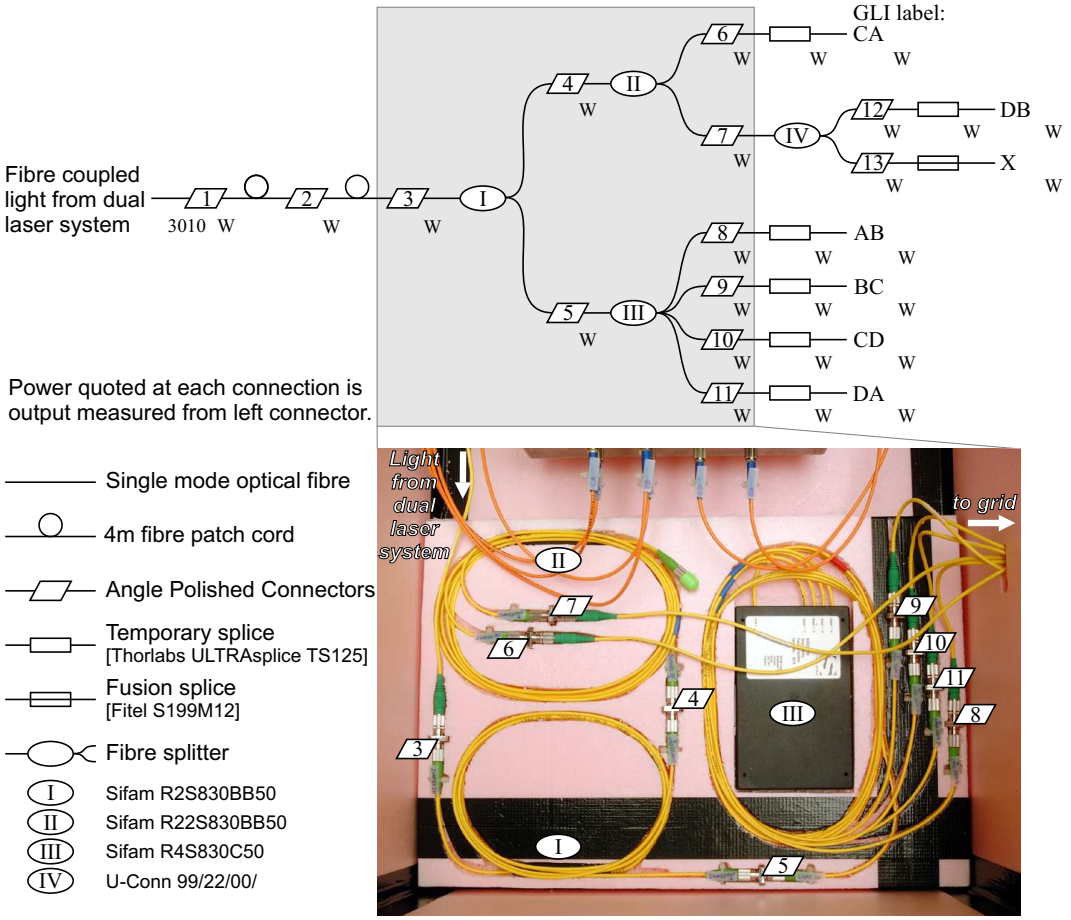

Figure 3.2: *The layout of the fibre splitter-tree that conveyed light from the dual laser system to multiple grid line interferometers. The optical power measured at each connection of the splitter-tree is shown.*

The fibre splitter-tree initially had 6 outputs connected via temporary splices to the delivery fibres of the 6 GLIs. In this set-up, the power of each long GLI was approximately twice that of each short GLI. When a seventh interferometer was later added (for reasons discussed in Section 3.3.3), the power was taken from the long GLI with the strongest return signal. The seventh

interferometer was connected by fusion splicing, due to the availability of suitable equipment. The temporary splices passed sufficient light to not require reconnection by fusion splicing.

All connectors in the splitter-tree were angle polished to suppress internal reflections that would otherwise interfere with and degrade the GLI signals. No significant GLI signal degradation due to splitter-tree reflections was observed.

3.2.3 GLI Readout with APDs

The power of the beam emerging from each quill was typically 50 μW and the return power was of order 10 pW.

The GLI signals were read out by Hamamatsu S2383-01 avalanche photodiodes embedded in a custom built multichannel board [Mit02], shown in Figure 3.3. This device was completed with seven further FC to ferrule optical patch cords that were mated with APD-fibre couplers on the board, enabling simultaneous measurement of up to eight channels.

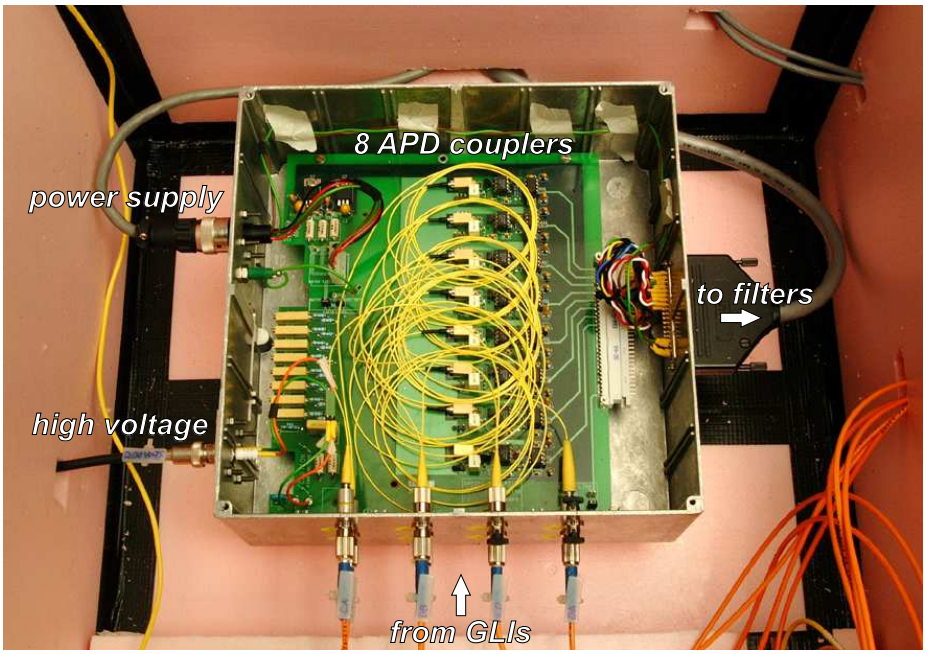

Figure 3.3: *The avalanche photodiodes that read out the GLI signals.*

The bias settings of the APDs were retained from previous work [Mit02].[2] The voltage applied across the bias chain was typically -180 V during the FSI

[2]Only channels 0-6 were used. Due to a negative output voltage for channel 5, the amplifier microchip at channel 5 was exchanged with that at channel 7.

measurements described in Chapter 4. The APD gain was slightly temperature dependent [Mit02] so the APDs were thermally insulated to slow down APD gain changes during FSI measurements.

The mean voltage of a GLI signal recorded by the ADC was typically $V_{signal} = 2$ V. The corresponding mean power incident on the APD was [Mit02, Mit03],

$$\begin{aligned} P_{signal} &= \frac{V_{signal}}{\mathcal{R}MR_FAF} \\ &\approx 1\ pW \end{aligned} \tag{3.1}$$

where:
Responsivity, $\mathcal{R} \approx 0.5$ AW^{-1},
APD gain, $M \approx 100$ (for -180 V bias at 22°C),
Feedback resistor, $R_F \approx 82$ MΩ,
Voltage gain, $A \approx 40$,
Filter-amplifier gain, $F \approx 12.5$.

For an ideal quill delivering $P_{delivered} = 50$ μW towards a perfect retro-reflector 0.2 m away, the power of light coupled into the return fibre of the quill from the retro-reflector (R-mode) should be approximately [Mar77],

$$\begin{aligned} P_R &\approx P_{delivered}\left(\frac{2\pi {w_0}^2}{\lambda D}\right)^2 \\ &\approx 2\ pW \end{aligned} \tag{3.2}$$

where:
wavelength, $\lambda = 830$ nm,
double GLI length, $D = 0.4$ m,
waist of the Gaussian beam, $w_0 = 3.4$ μm, calculated from the fibre specifications [Fibre] at $\lambda = 830$ nm.

The typical visibility of the FSI fringes, $\gamma \approx 2/3$, so from Equation 2.8, assuming $P_V > P_R$,

$$P_V \approx 6.9 P_R. \tag{3.3}$$

Combining Equations 2.6, 3.2 and 3.3,

$$P_{DC} \approx 16\ pW. \tag{3.4}$$

After accounting for fibre connection losses, the measured mean power of the signal, P_{signal} was therefore compatible with the expected return power due to the divergence of the Gaussian beam emerging from the quill given the length of the GLI.

3.3 Prototype Grid Design and Construction

3.3.1 Overview

Two prototype grids were constructed by different arrangements of the same grid components. Each grid consisted of six distance measurements between four nodes. The first grid had the nodes placed at the corners of a square so that the measurement lines formed a braced-quadrilateral[3], as shown in Figure 3.4. This over-constrained geometry allowed the reconstruction process to be checked for self-consistency. The second grid had a tetrahedral geometry, described below. Each grid had one node mounted on a motion stage to permit fine adjustment of the grid shape.

3.3.2 FSI Jewels

Ideally the distance measurements would have emanated from point like grid nodes. In practice each grid node was defined within an extended object called a *jewel*: a small stable support structure that accurately positions a cluster of quills[4] and/or retro-reflectors. A quill from one jewel pointed towards a retro-reflector in another jewel to form a grid line interferometer (GLI). The quill and retro-reflector separation was measured by FSI.

In principle, one retro-reflector may be common to several GLIs that converge within the angular acceptance of the retro-reflector. In practice, the studied grid geometries required multiple retro-reflectors in some jewels, due to the large angular separation ($\sim 45^\circ$) of convergent grid lines. The jewel that defined node B of the grid supported two retro-reflectors and one quill, as in Figure 3.5. The GLI components were held in 6 mm diameter spheres positioned at a constant radius from the grid node, where the extended lines of the GLIs converged. The distance measurements were therefore orthogonal to movements of the GLI components caused by small jewel rotations about the grid node, reducing systematic errors by design. The front surface of the jewel was angled down by 4° to suppress spurious reflections.

3.3.3 Variable Grid Geometry

The 4 grid nodes are labelled A-D. 3 nodes were fixed but node C was mounted on a motion stage (Melles Griot 17AMU 001/D), as in Figure 3.6. Node C could be manipulated in the two translation and two rotation axes of the stage. Each axis had a fine thread and UltraFine^{TM}differential drive, although neither the

[3]A braced-quadrilateral is the simplest over-constrained figure used in surveying.

[4]described in Section 2.2.2

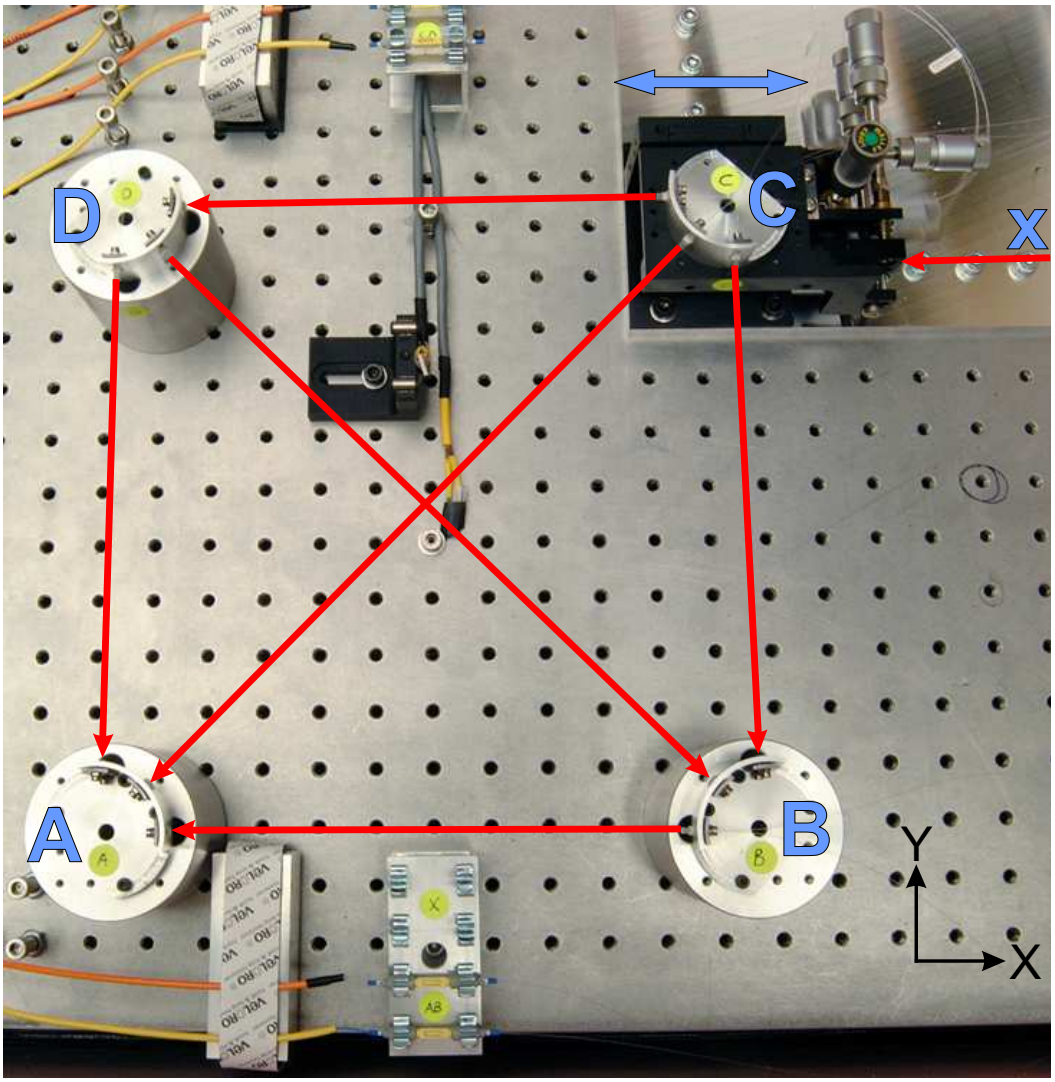

Figure 3.4: *A prototype FSI grid formed by 6 grid line interferometers between the corners of a square. The 6 distance measurements formed a braced-quadrilateral. A seventh interferometer provided an independent measurement of the node C stage displacement. The grid was 254 mm by 254 mm square.*

rotation axes nor the fine translation threads were adjusted after the position and orientation of jewel C were initially set. Node C was only displaced with the differential translation drives, between FSI measurements. No scale was present on these drives, so to read and set the displacement of node C repeatably, a fine wire was secured next to each drive adjustment arm, against which a mark on the arm was compared. Despite the claimed 50 nm setting sensitivity, the

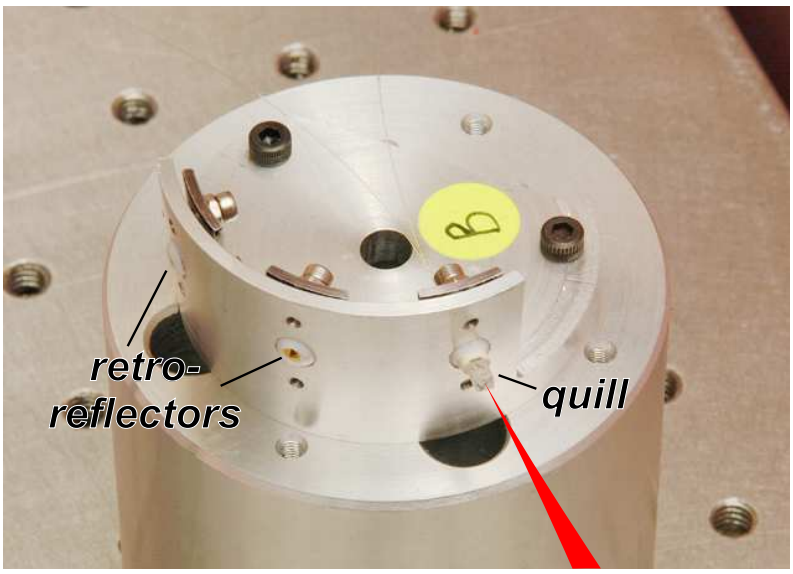

Figure 3.5: *The jewel that defined grid node B was mounted on a pillar bolted to the optical breadboard. The jewel held two retro-reflectors and one quill for the three GLIs convergent on the node. The GLI components were held in 6mm diameter spheres, clamped into the circular support frame of the jewel.*

resolution on reading and setting the stage displacement was around ± 1 μm due to the thickness of the mark and wire and small parallax errors.

The resolution on reconstructing the stage displacement from FSI grid measurements was finer than could be achieved by a manual reading of the stage displacement. Therefore a seventh FSI interferometer, labelled X, was introduced to independently measure the stage displacement. Further interferometers to measure the orthogonal stage displacements were anticipated in designing the apparatus, but were never realized due to lack of the fibre-splitters required for additional simultaneous measurements.

The grid was later adapted to the tetrahedral geometry shown in Figure 3.7. Node C was raised 100.0 mm by inserting four accurately machined pillars under the aluminium plate that carried the motion stage and the seventh interferometer. The resulting out-of-plane component of the GLIs convergent on node C allowed three dimensional reconstruction of the nodes to be investigated.

3.3.4 GLI components

The quills in the prototype grids were manufactured by the same method as those installed in the ATLAS FSI grid, with one exception: the beam-splitters were glued to the quill bodies using an adjustable prototype jig, rather than the

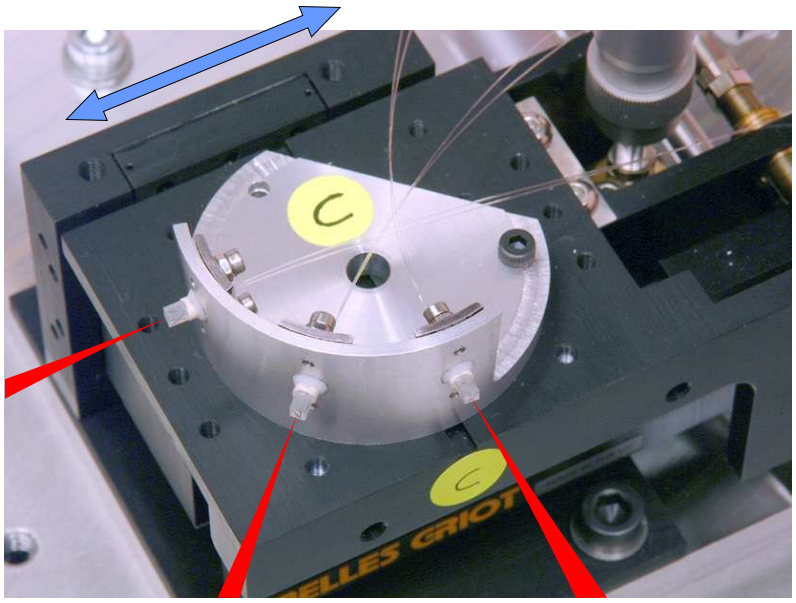

Figure 3.6: *Grid node C was mounted on a motion stage to permit fine adjustment of the grid shape. The glass beam-splitters of three quills can be seen protruding from the 26 mm radius circular support frame. The delivery and return fibres for these quills trail from the back of this frame.*

jig later designed for mass production. The beam-splitters used were 2.2 mm by 2.2 mm square with a thickness of 2.9 mm from the quill body to the farthest edge. The front surface of each beam-splitter was angled by approximately 4°. The quills were selected if the two fibres within each quill were parallel to less than 0.5° [Reic01]. The quill fibre had an optical specification [Fibre] matching that of the FSI fibre for ATLAS, but was not radiation hard.

The retro-reflectors were punched from aluminium slugs with a carbon steel corner cube, cleaned and then coated with nickel followed by gold to enhance reflectivity [Huf01]. Due to variable production quality, suitable retro-reflectors for the prototype grids were selected, as in Section 3.5.1.

3.4 Alignment of Grid Components

Jewels A, B and D were mounted on pillars bolted to the optical breadboard. Newport scientific grade breadboards have a surface flatness specification of ±0.1 mm over 600 mm square [New03]. In combination with the machining tolerance of the pillars, this ensured the centre of these 3 jewels were planar to within ±0.2 mm.

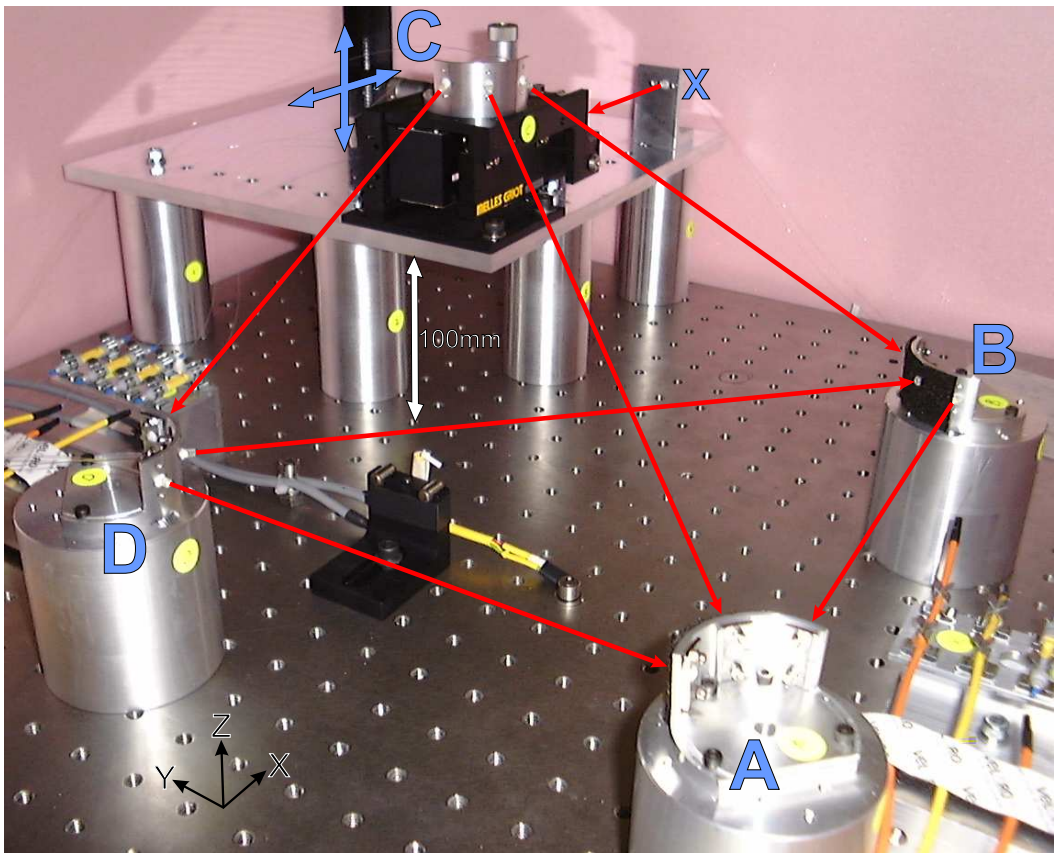

Figure 3.7: *A three dimensional prototype FSI grid with tetrahedral geometry. Grid node C was elevated 100 mm above the plane ABD.*

Jewel C was mounted on the motion stage and the rotation axes were set level and square. Before the GLI components were inserted into the jewels, a fine wire was stretched between the quill and retro-reflector position of each GLI. The vertical axis of the motion stage was adjusted until the two diagonal wires that crossed in the centre of the grid just touched, ensuring the planarity of all four nodes to ±1 mm. The planarity and 254 mm separation of the grid nodes was confirmed with ruler measurements to $< \pm 1$ mm.

The GLI components were held in 6 mm diameter spheres to allow pointing adjustment during initial alignment. Each retro-reflector sphere was adjusted so that its axis was aligned with the relevant quill. No further adjustments were necessary due to the large angular acceptance of the retro-reflector of around ±10°. Each quill-sphere was adjusted as shown in Figure 3.8 with Laser 1 on full power. The angular divergence of the Gaussian beam emerging from each quill was ±4.5° (for the $1/e^2$ irradiance contour) so the diameter of the beam at

the retro-reflector was about 30 mm for short lines and 45 mm for the long lines, much larger than the Ø2.5 mm retro-reflector. The centre of the Gaussian beam was aimed at the retro-reflector to maximize the reflected signal. The reflected signal and the partial reflection from the quill beam-splitter were coupled into the return fibre to a remote APD, which measured the resulting interference. As the frequency of laser light was scanned, sinusoidal interference fringes were monitored on an oscilloscope. Further fine adjustments to the pointing of the quill were made to maximize the visibility of this sinusoidal interference. ATLAS FSI jewels are pre-aligned in a jig by the same method, as in Section 2.8.

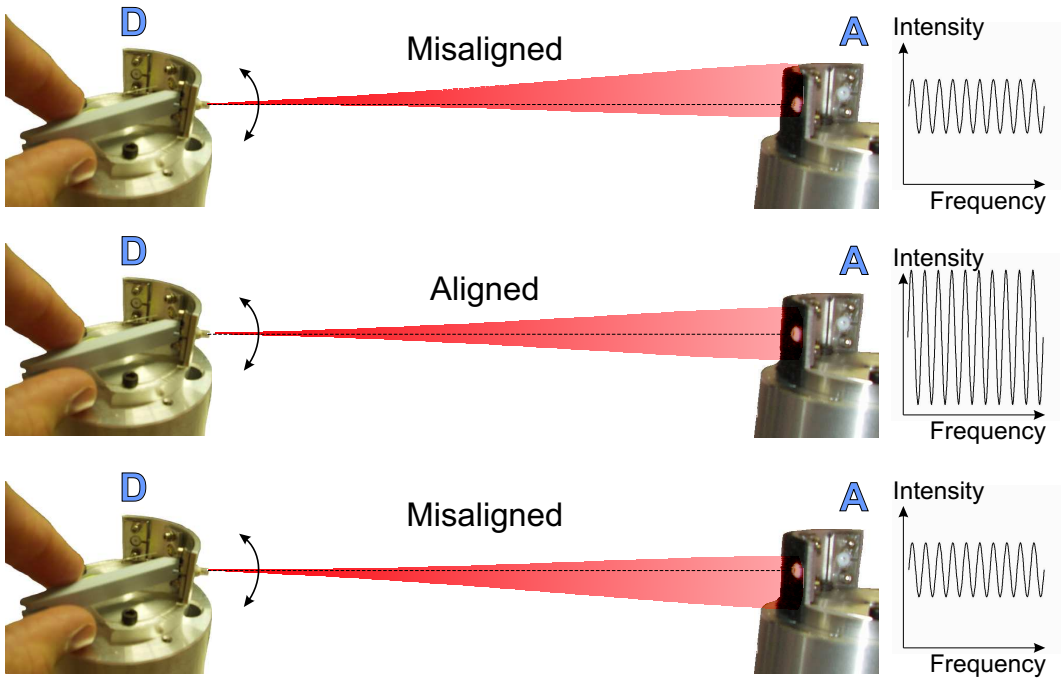

Figure 3.8: *Pre-alignment of a grid line interferometer. A quill in jewel D targeted a retro-reflector in jewel A. A tool inserted into two pin holes through the quill-sphere allowed the quill pointing to be adjusted to maximize the GLI signal visibility.*

3.5 GLI Signal Optimization

The majority of GLI signals were initially found to be of insufficient quality for reliable and precise FSI measurements. Various methods were employed to improve the GLI signals as described in this section.

3.5.1 Replacement Retro-reflectors

The grid was initially equipped with retro-reflectors randomly chosen from a batch of 24. The GLI signal quality from 5 of the 6 installed retro-reflectors, was too poor to reliably extract the GLI lengths during the analysis. Replacement retro-reflectors from the same batch reduced the spread on the GLI length predicted from each individual FSI sub-scan, as shown for one GLI in Figure 3.9. Before the retro-reflector was replaced, the large spread adversely affected linking the GLI phase. The linking extrapolation from the average mean q_0 of the FSI fine tuning sub-scans did not lie within the required $\pm\pi$ GLI radians of the drift corrected target value [Coe01]. The linked GLI length was therefore severely degraded, as in Figure 3.9. On replacing the retro-reflector, the spread reduced sufficiently for the remaining errors to be significantly reduced by linking.

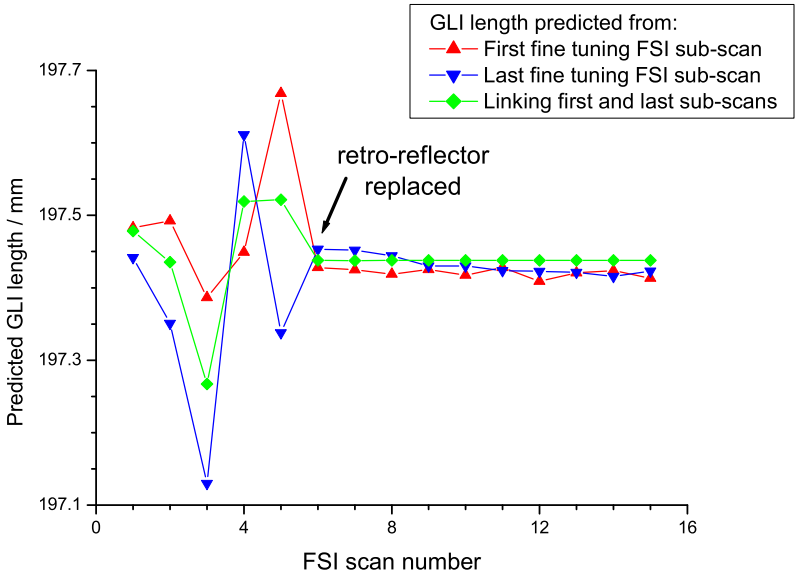

Figure 3.9: *The effect of changing a retro-reflector on the predicted GLI length for both FSI sub-scans and after linking both sub-scans. On replacing the retro-reflector the spread on the prediction from each sub-scan reduced sufficiently for successful linking (FSI scan 12 required one GLI order correction).*

Despite the replacement retro-reflector, FSI scan 12 required one GLI order correction to the predicted GLI length. This was due to the large 3.5 THz single linking bridge inducing an extrapolation error outside of the required $\pm\pi$ GLI radians. The technique of multiple linking, described in Chapter 2, eliminated the need for such manual GLI order corrections and automated the linking process.

The above observations stimulated a search for retro-reflectors of an appropriate quality for FSI. Measurements with the WHIPM system [Huf01] allowed a set of 17 *very good* retro-reflectors to be selected from a batch of 192, on the basis of reflectivity and interferogram figure [Huf02]. Five of these were installed in the prototype grid and one formed part of the stage measurement interferometer. These retro-reflectors were found to be suitable for FSI measurements. One suitable retro from the original batch was retained. A rigorous quality assurance programme has been developed and implemented for retro-reflectors installed in ATLAS, as discussed in Section 2.7 [Coe02a, Huf01].

3.5.2 Introduction of Filter-amplifier

To improve the signal to noise, a multichannel 4-pole Butterworth low-pass filter-amplifier [Hol66] was introduced between the APDs and ADC. The data acquisition rate and frequency tuning rate were retained from previous work [Mit02]. The time response of the filter was chosen so that any ringing had settled to 1 bit of the 12 bit ADC just before the APDs were recorded in the last DAQ step of each chopper cycle. This maximized the smoothing effect of the filter without introducing a memory effect from the previous laser chopper cycle that might otherwise cause systematic errors. Initially only GLI data from the last DAQ step of each chopper cycle were collated for fitting in the analysis.

For the data presented in Chapter 4, the FSI measurement precision was further improved by around 10% by averaging the GLI intensity measurements from the last three steps of each chopper cycle. The first step was still excluded from the analysis. In previous measurements without the additional filter, the GLI data from four steps of each chopper cycle were averaged [Mit02].

A GLI signal was simultaneously recorded without and with the filter-amplifier, as in Figure 3.10. The best fit sinusoid amplitude I_{AC} was less for the unfiltered signal because the χ^2 could be lowered during the search for a minimum, by decreasing the value of I_{AC} [Coe01]. A significant improvement was found on the fit to the filtered data.

3.5.3 Suppression of Spurious Reflections

Optical cross-talk

Optical cross-talk was found between GLIs that illuminated the same grid node. The cross-talk was manifest as high frequency jitter in the GLI signal, shown in Figure 3.11. By temporarily obstructing light from the GLI at 45° incidence to the GLI under study, the cross-talk vanished. This indicated spurious re-

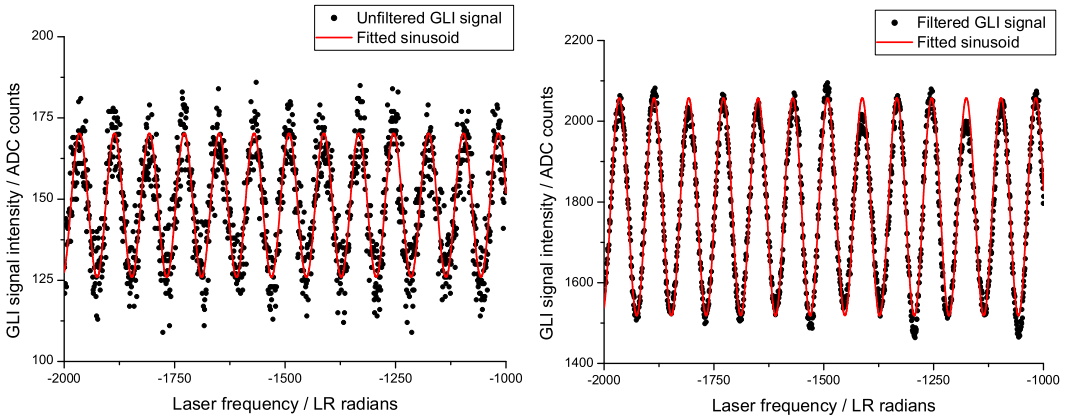

Figure 3.10: *A GLI signal simultaneously recorded without and with the filter-amplifier. A significant improvement was found on the fit to the filtered data.*

flections from the front curved surface of the jewel coupled light from one quill into the return fibre of another quill. The optical cross-talk did not significantly degrade the fitted GLI signal χ^2, as shown in Figure 3.11. The reflections were suppressed by adding a diffuse black material to the front surface of jewels A and B, where this problem was apparent.

Potential spurious reflections that couple light between GLIs were avoided by design in the ATLAS FSI grid, as discussed in Section 2.8. Should any severe optical cross-talk become apparent in the ATLAS FSI grid, the problem could be eliminated by staggering the illumination of multiplexed groups of GLIs, that do not exhibit optical cross-talk.

Direct reflections

Some GLI signals were also degraded by spurious reflections from the retro-reflector adjacent to that of a given GLI, even though the axis of the adjacent retro-reflector was angled away by 45°. In these cases, the axial orientation of the adjacent retro-reflector was such that one of its three mirrored planes was almost perpendicular[5] to the incident light of the given GLI. The spurious reflection coupled light back to the same quill and interfered with the reflection from the correct retro-reflector. As before, by temporarily obstructing the spurious reflection, the signal degradation vanished. The spurious reflection was eliminated by rotating the adjacent retro-reflector about its axis, until none of the three planes of that retro-reflector were close to perpendicular to the incident ray of the GLI that exhibited a degraded signal.

[5]within 7°

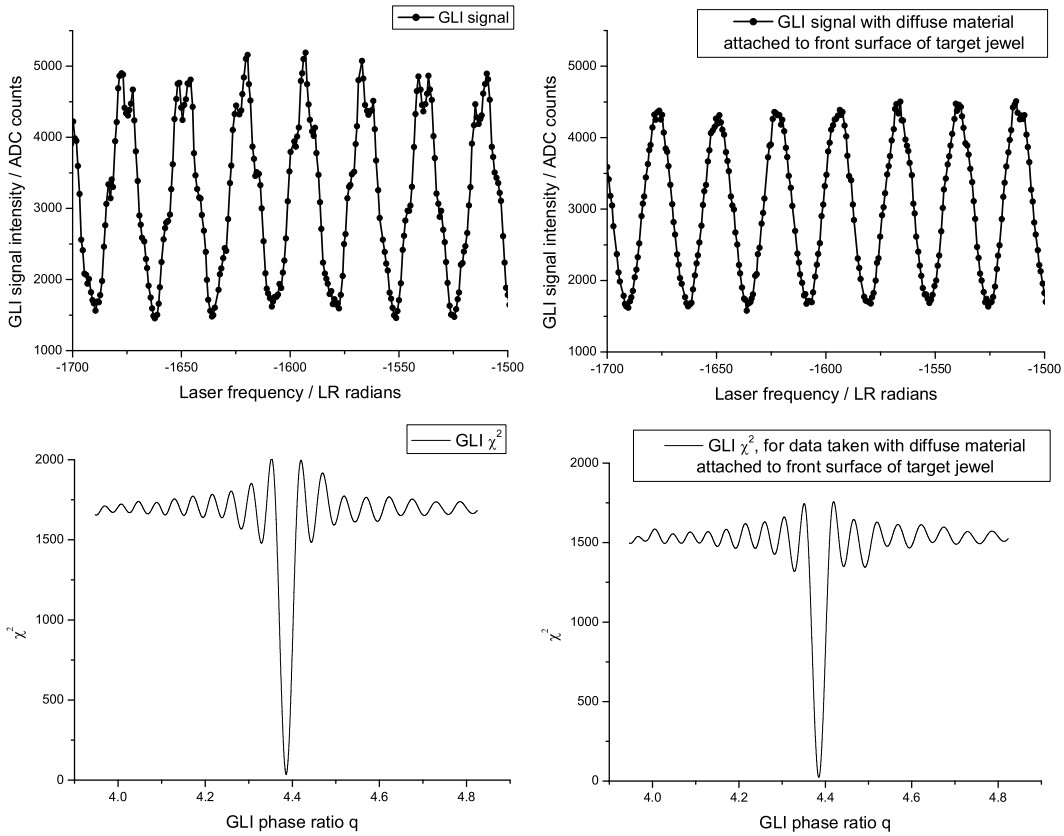

Figure 3.11: *The degradation of a GLI signal caused by optical cross-talk from an adjacent GLI, incident on the same target jewel. The cross-talk vanished when the adjacent GLI was temporarily obstructed. Adding a diffuse material to the front surface of the target jewel reduced the cross-talk present in the GLI signal. In either case, the GLI χ^2 profile exhibited a distinct, deep minimum, into which the fit settled.*

This type of spurious reflection should be avoided in the ATLAS FSI grid by design, as discussed in Section 2.8.

3.5.4 Further Improvements

The following methods also improved the recorded GLI signals:

- Adjustment of the two temporary splices of each GLI.
- Adjustment of the optical connectors within the fibre splitter-tree and between the temporary splices and the APDs.

- Replacement of one quill that had a significant W-mode reflection.
- Observance of silence during FSI measurements, to avoid undue vibration of the set-up. The tolerance of the ATLAS FSI system to SCT vibrations is discussed in Section 2.6.2 [Coe02b].

3.6 Conclusion

The first, prototype FSI geodetic grids have been constructed as part of a prototype ATLAS FSI system that has demonstrated remote, multiple, simultaneous, distance measurements between grid nodes.

The grid was powered via an optical fibre splitter-tree, distributing about 50 μW to each grid line interferometer. The 10 pW return signals were successfully read out with multiple APDs.

Two grid geometries have been constructed by different arrangements of the same grid components. In each grid, the position of one grid node was finely adjustable. The grid components reduced by design systematic errors induced by spurious reflections and small jewel rotations.

The GLI components were of the same design as for ATLAS. Each GLI was aligned to maximize the interference signal during a frequency scan. The GLIs signals were improved by various methods including: replacement of poor quality retro-reflectors; introduction of a filter-amplifier; suppression of spurious reflections. The following should be noted for future FSI work:

- Optical cross-talk was observed between GLIs incident on the same target jewel. This cross-talk did not significantly degrade the fit to the GLI signals, but spurious reflections of this type should be avoided in FSI jewel design where possible.
- Light incident from outside the nominal angular acceptance of a retro-reflector may be returned, via only one or two of the mirrored surfaces. A retro-reflector should be orientated about its axis to avoid returning light to any quill not intended to target that retro-reflector.
- The majority of retro-reflectors from a manufacturing batch were found to be unsuitable for precise FSI measurements. The retro-reflectors installed in ATLAS should be appropriately evaluated and selected.

The FSI geodetic grids were set up to enable the tests described in Chapter 4.

Chapter 4

FSI Grid Shape Measurement

4.1 Introduction

ATLAS SCT shape deformations will be reconstructed from FSI data by combining the distance measurements between nodes attached to the SCT. The distance measurements form a geodetic grid that will be solved for the node positions and hence the grid shape. Further software will then interpolate for the module co-ordinates as discussed in Section 5.5.

This chapter concerns the first demonstration of FSI geodetic grid reconstruction, with measurements of the prototype grids described in Chapter 3. The prototype grids were configured for various tests of the measurement and reconstruction method. The primary aim was to demonstrate prototype grid reconstruction within the precision required for ATLAS.

To determine a particular grid shape, the distances between grid nodes were simultaneously measured with the prototype FSI system. A software model then combined the FSI measurements to reconstruct the grid node co-ordinates. The software model also predicted the precisions on reconstructing the grid node co-ordinates, which were evaluated with repeated FSI grid measurements. To validate the grid reconstruction method, the data were checked for:

- compatibility of the predicted and measured precisions on reconstructing the grid node co-ordinates.
- self-consistency in reconstructing an over-constrained grid.
- linearity in reconstructing the displacement of grid nodes.

The tests also sought to identify any issues related to the measurement and reconstruction of the ATLAS FSI grids, and how these issues might be addressed.

4.2 Grid Shape Measurement Method

4.2.1 General Method

The FSI grid measurement and reconstruction method outlined in this section, was tested in various ways with the prototype grids. In each test, the grid shape was repeatedly measured and reconstructed, to assess the precision in determining the grid node co-ordinates. In certain tests the grid shape was altered between measurements by adjusting the position of the grid node mounted on the motion stage. The differences between the reconstructed grid shapes were compared with the induced grid shape changes. In most tests, at least 5 grid measurements were performed at any given stage position so that measurement errors could be assessed.

4.2.2 FSI Grid Measurements

The first step in determining the grid shape was an FSI scan that simultaneously measured all GLIs on the time scale of a few minutes. The data presented in this chapter were acquired by the FSI *multiple linking* technique, detailed in Section 2.4.4, in which a set of fine tuning FSI sub-scans are interspersed with periods of coarse tuning, to extend the effective frequency range of the scan and improve the measurement precision.

The fine tuning FSI sub-scans of Laser 1 each had a frequency range of about 60 GHz. The combined frequency range of the coarse tuning intervals was about 6.9 THz, in all tests. The total linked frequency range of Laser 1 was centred on the midpoint frequency of Laser 2 (midpoint wavelength $\approx$ 835.79 nm). The coarse tuning pattern that defines the relative frequency separation of the fine tuning sub-scans, was chosen so that the phase extrapolation for successive links was within the required $\pm\pi$, with the minimum number of links, to reduce total measurement time and hence the adverse effects of interferometer drift.

In some frequency regions, Laser 1 was observed to exhibit brief bursts of multi-moded lasing during frequency tuning, which considerably reduced the visibility of all interferometer signals and sometimes introduced a phase unwrapping glitch that was too severe to be automatically identified and corrected by the algorithm described in Section 2.3.3. In these regions, the output of an optical spectrum analyser showed between three and six Fabry-Pérot modes lasing simultaneously. The coarse tuning pattern was therefore adjusted to avoid the fine tuning intervals coinciding with frequency ranges of Laser 1 that were prone to multi-moded behaviour. The course tuning pattern set when acquiring each data set is shown in Table 4.1.

Data set			Coarse Tuning Pattern [nm]	Equivalent Frequency Intervals [THz]
A	square	linearity	1.2 10.0 4.8	0.5 4.3 2.1
B	tetrahedral	precision	1.2 10.0 4.8	0.5 4.3 2.1
C	tetrahedral	linearity	4.5 10.0 1.5	1.9 4.3 0.6
D	tetrahedral	planarity	4.5 10.0 1.5	1.9 4.3 0.6
E	tetrahedral	FSI	4.5 10.0 1.5	1.9 4.3 0.6

Table 4.1: *The coarse tuning patterns used for each data set. The actual frequency intervals were dependent on the slightly variable coarse tuning speed of Laser 1, which may have reduced the largest link by up to 1 nm (0.4 THz).*

4.2.3 Grid Shape Reconstruction

The grid shape was reconstructed for each FSI scan by following the analysis procedure summarized in the flow diagram of Figure 4.1.

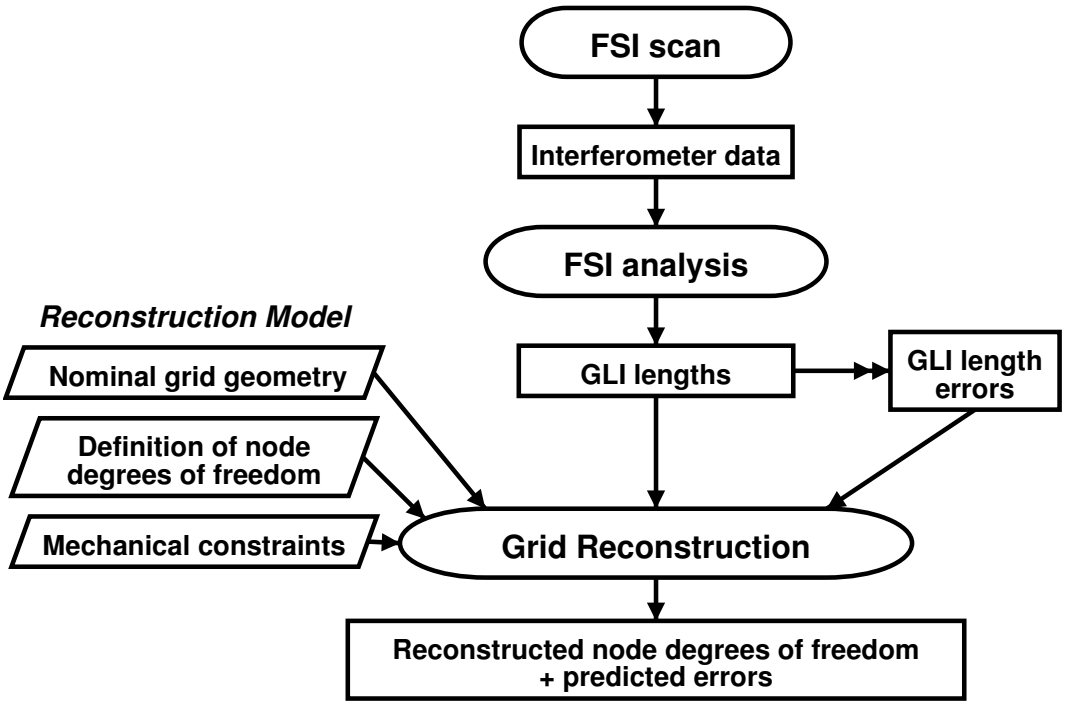

Figure 4.1: *Overview of analysis procedure to reconstruct the prototype grid shape from FSI measurements.*

The GLI lengths were extracted from the interferometer data by the techniques detailed in Chapter 2. Additionally, the following information and definitions were required to reconstruct the grid shape:

- nominal grid geometry and jewel design.
- a definition of co-ordinate frame in which to reconstruct the grid nodes.
- a definition of degrees of freedom of the grid nodes.
- the mechanical constraints.
- the scale.
- precisions on each FSI length measurement.

This information was coded in a model of the grid, written in the SIMULGEO software package [Bru98]. The model was used to combine the GLI length measurements from a given FSI scan to reconstruct the grid node co-ordinates.

4.2.4 SIMULGEO Reconstruction Software

SIMULGEO [Bru98] is a general purpose tool for the simulation and reconstruction of opto-geometrical systems. The software was developed between 1995-2000 in collaboration with the CERN surveying group and has found application in various HEP projects including CMS, ALICE and the NLC alignment and survey (LiCAS). The following features commend the software to the simulation and reconstruction of FSI geodetic grids:

- a three dimensional geodetic grid of distance measurements between nodes may be readily defined.
- the distance measurements need not converge on point-like grid nodes; rather, distinct measurement nodes corresponding to GLI components may be embedded within a co-ordinate frame to define an FSI jewel.
- the 6 external degrees of freedom (DoF) of a jewel that define its position and orientation, may be specified and if required, independently constrained.
- the internal DoF of a jewel, such as the separation of GLI components within the jewel, may be similarly specified and constrained.

- by imposing constraints, mechanical properties may be modelled. The constraints can be understood as the resistance of the constrained parameters to be adjusted during the reconstruction.
- the precision of each FSI measurement is accounted for in the adjustment procedure.
- redundancy loops are accounted for.
- the algorithm has been optimized for calculation time and memory consumption.
- a graphical user interface and display allows quick visualization of the model.
- the software is written in the java language, has object-oriented design and is platform independent.

Modes of operation

SIMULGEO has two modes of operation in the context of FSI geodetic grids:

error propagation: the precisions on the free parameters in the model (usually the grid node co-ordinates) are calculated from the precisions of each GLI length measurement, the nominal grid geometry and the constraints.

reconstruction: the best estimate of the free parameters (usually the grid node co-ordinates) are computed from a set of GLI length measurements and precisions. The computation also requires the nominal grid geometry and constraints.

Assumptions of algorithm

The algorithm encoded in SIMULGEO computes the error propagation by the variance-covariance method and the reconstruction by an iterative least-squares method [Bru98]. In both cases the algorithm assumes that the parameter errors are Gaussian and uncorrelated.

4.3 Square Grid

4.3.1 Overview

This section describes studies of grid reconstruction with FSI measurements of the square prototype grid described in Section 3.3. Initially, a model of the

square grid was defined and an error propagation was performed with SIMULGEO, to predict the precisions on reconstructing the node co-ordinates. These predictions were checked with Monte Carlo data, to ensure the reconstruction proceeded as expected.

The grid shape was repeatedly measured with FSI, following the procedure summarized in Section 4.3.4. The extracted GLI lengths were checked for correlations with the temperature of the optical bench, as in Section 4.3.5.

Reconstructing the grid shape from FSI measurements, required calibrating the model, as summarized in Section 4.3.6. The redundancy in the grid allowed the self-consistency of the reconstruction to be checked and compared with prediction from simulation, as in Section 4.3.8. The linearity of the reconstruction process was also examined, as in Section 4.3.9.

4.3.2 Reconstruction Model Definition

The square grid was reconstructed with a SIMULGEO model in which the DoF of the grid components were defined as follows. Node A defined the origin, the line from node A to node B defined the X-axis and all four nodes were assumed to be coplanar. The rotational DoF of the jewels were assumed to be fixed. The distance between nodes A and B defined one DoF and nodes C and D were free to translate in the X-Y plane but constrained in Z, giving 5 DoF in total. The 6 distance measurements thus provided an overconstrained geodetic grid.

For the initial error propagation described in the next section, perfect geometries of the jewels were specified in the model from engineering drawings. The jewel geometries were later calibrated when combining real FSI measurements as described in Section 4.3.6. The separation of GLI components within a jewel was assumed constant after the calibration.

4.3.3 Preliminary Reconstruction Studies

Error Propagation

An error propagation was performed with the SIMULGEO model to predict the precision on reconstructing the grid node co-ordinates. A precision of $\sigma = 1$ μm was assumed for all length measurements. The predicted precisions are shown in Figure 4.2. Node A defined the origin so has no error. Node B defined the X-axis so has no error in the Y co-ordinate. The predicted precision on certain node co-ordinates was less than the assumed measurement precision of 1 μm due to redundancy.

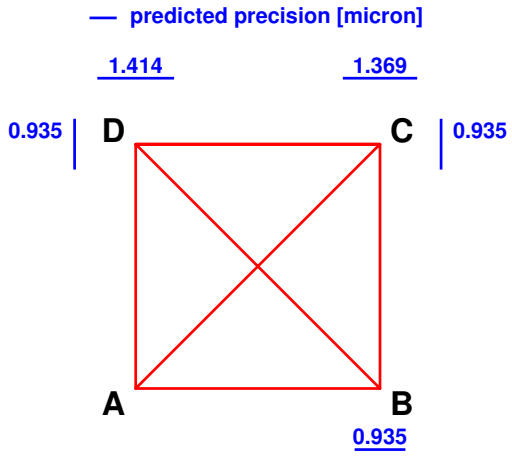

Figure 4.2: *The predicted precisions on the five free parameters of the square grid, assuming a 1 μm FSI measurement precision. The line lengths and numbers indicate the predicted precision on reconstructing the co-ordinate of the adjacent grid node, in the line direction.*

Check with Monte Carlo

A set of simulated length measurements were reconstructed with the SIMULGEO model, as a basic check of the reconstruction process. An independent model of the square grid was written (not in SIMULGEO), to describe a perfect geometry, in which the grid nodes were separated by exactly 254 mm. The GLI lengths were taken from this model and 10,000 copies of each length were independently smeared with a Gaussian of width 1 μm to simulate length measurement errors. The grid node co-ordinates were then reconstructed in SIMULGEO from each set of smeared GLI lengths and are plotted in Figure 4.3.

Importantly, the mean values of the reconstructed co-ordinates matched those in the original model geometry. The precision on each reconstructed grid node co-ordinate was determined by a Gaussian fit, as in Figure 4.4. The reconstructed precisions are compared with those predicted from the error propagation in Table 4.2. The precisions are compatible within the statistical errors estimated by Minuit [Jam75], showing that in this simple check, the reconstruction proceeded as expected.

The Monte Carlo check also revealed the correlations between certain free parameters, such as the reconstructed X and Y co-ordinates of node C. The major axis of the elliptical scatter of node C is not parallel to the X or Y axes, primarily due to the additional diagonal line CA, reducing the scatter in the

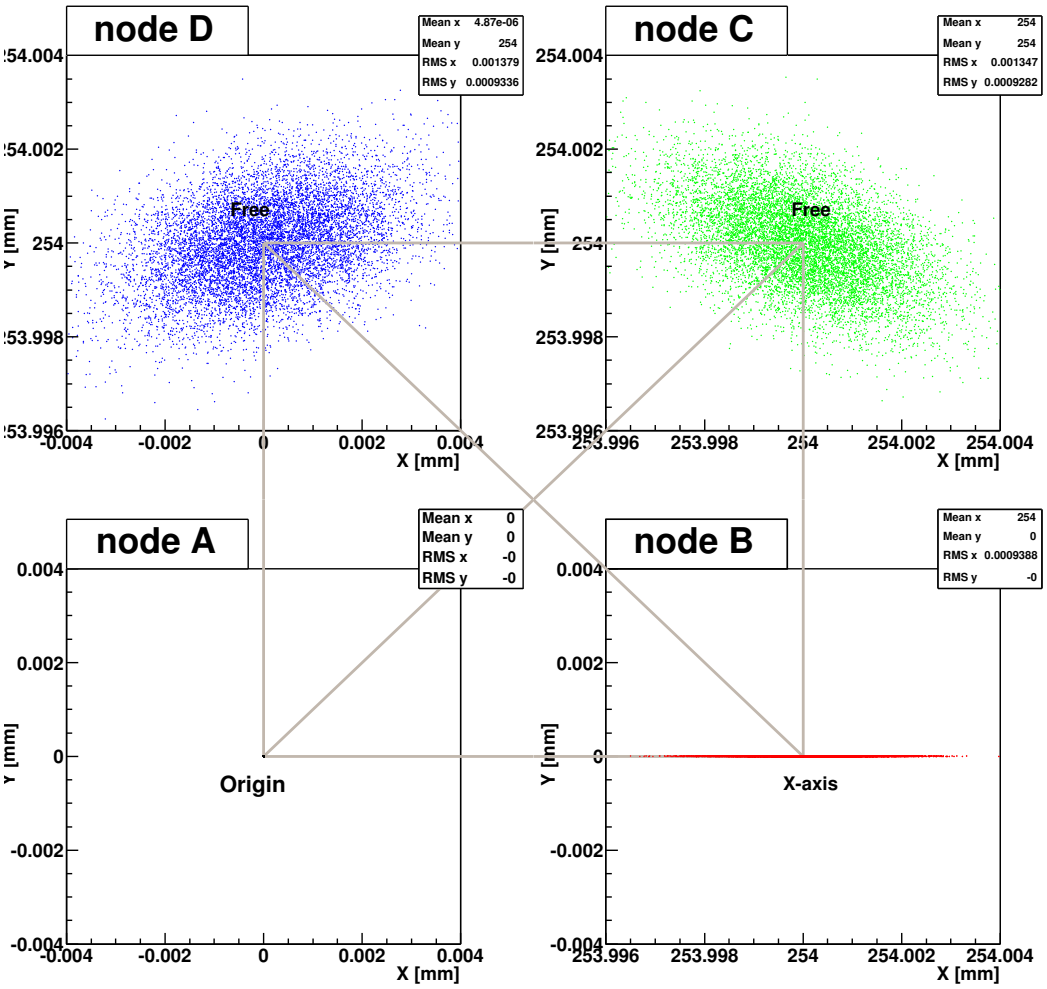

Figure 4.3: *The simulated scatter in reconstructed co-ordinates of the four nodes of the square grid, due to a 1 µm smearing of the GLI lengths, representing the required ATLAS FSI measurement precision. The plot emphasizes the impact of the DoF definitions in the model, in which node A defined the origin in the model and node B defined was constrained to lie along the X-axis. Histograms of the scatters are shown in Figure 4.4.*

$Y = X$ direction. The asymmetric DoF definitions of the model also affect the correlation. The effects of similar correlations should be considered when defining an appropriate set of DoF for the ATLAS FSI grid.

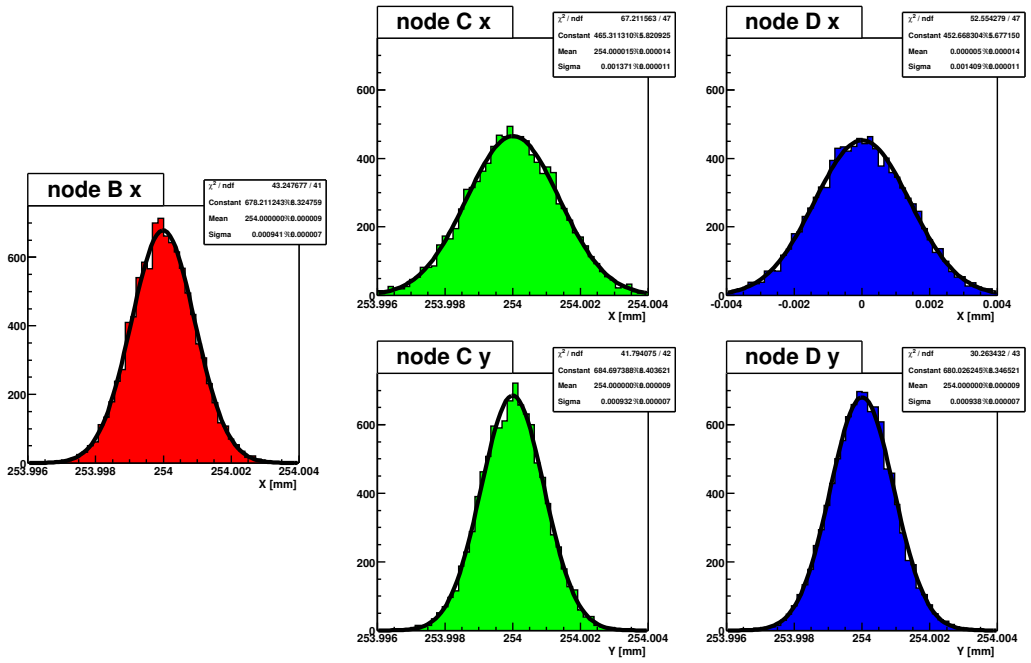

Figure 4.4: *Gaussian fits to the simulated scatters in reconstructed co-ordinates of the square grid nodes, due to a 1 μm smearing of the GLI lengths.*

Grid node co-ordinate	Predicted precision [μm]	Simulated reconstructed precision [μm]	Fit error on reconstructed precision [μm]	Reconstructed minus predicted [μm]
Bx	0.935	0.941	$\pm$0.007	+0.006
Cx	1.369	1.371	$\pm$0.011	+0.002
Cy	0.935	0.932	$\pm$0.007	-0.003
Dx	1.414	1.409	$\pm$0.011	-0.005
Dy	0.935	0.938	$\pm$0.007	+0.003

Table 4.2: *Comparison of predicted and simulated reconstructed precisions on the square grid node co-ordinates, assuming a 1 μm measurement precision required for the ATLAS FSI system.*

4.3.4 Measurement Procedure

For the square grid, 35 FSI scans were recorded as node C was displaced along the X-axis of the translation stage. 5 FSI scans were made at each of the 7 stage positions. The differential translation drive was turned once between each set of five scans, defining the step size of 47 μm $\pm$1 μm.

The stage was then rewound and 5 further FSI scans were recorded at each

of the first 4 stage positions (within stage setting errors).

For each of the 55 FSI scans, the measured GLI lengths were extracted by the techniques of Chapter 2.

4.3.5 Consideration of Thermal Expansion

The temperature of the steel optical breadboard was monitored throughout the FSI scans so that the effects of thermal expansion on the measured GLI lengths could be studied. No significant correlation with the steel temperature was found for any GLI length. This was primarily due to the very small maximum change in steel temperature of $\Delta T = 0.087$ K observed during these tests. The expected corresponding length change for one of the short GLIs of length $L = 196$ mm is

$$\begin{aligned} \Delta L &\approx \alpha_{steel} L \Delta T \\ &\approx 250 \text{ nm} \end{aligned} \tag{4.1}$$

where the coefficient of expansion for stainless steel (18Cr+8Ni) $\alpha_{steel} = 14.7 \times 10^{-6}\ K^{-1}$ [Kay89]. The expected ΔL was of the same order as the FSI length measurement precision. Therefore any correlation of GLI length with the steel temperature could not be significantly resolved.

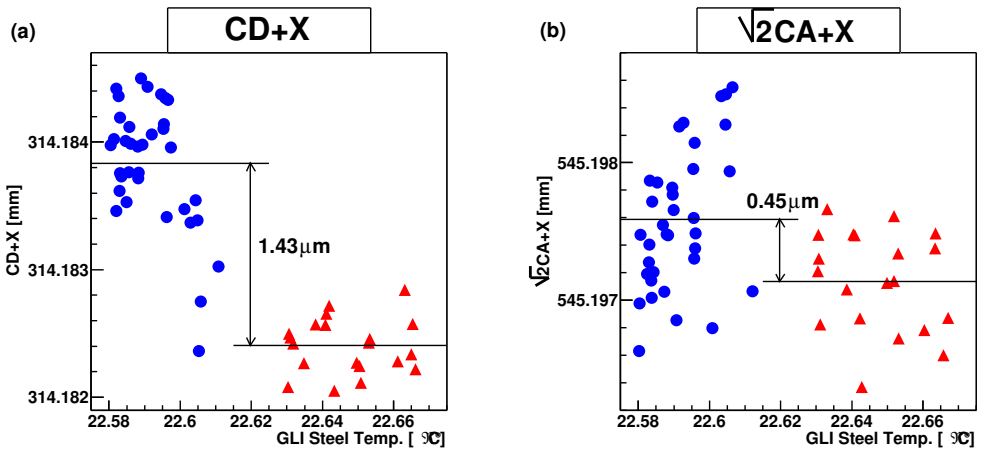

Figure 4.5: *The sum of lengths for interferometer X (measuring the stage displacement) and (a) GLI CD or (b) GLI CA×$\sqrt{2}$; versus the scan averaged steel optical breadboard temperature. These summed lengths should have been approximately independent of stage displacement. The blue circles and red triangles represent data before and after the stage was rewound. A significant discontinuity is present between the mean of the two data sets.*

During the FSI scans interferometer 'X' independently measured the displacement of the stage, on the opposite side of the stage to GLI CD. The measurements CD and X, and the stage translation were approximately collinear, so it was expected that the sum of lengths CD and X would remain constant, within measurement errors, for any stage translation. Instead, a significant discontinuity was observed when these summed lengths were plotted against the scan averaged GLI steel temperature, as shown in Figure 4.5(a). A discontinuity was also observed in a similar plot of the sum of X and GLI CA multiplied by $\sqrt{2}$ for the geometry, shown in Figure 4.5(b). In both cases the discontinuity occurred when the stage was rewound. The magnitude and sign of the discontinuity cannot be explained by the thermal expansion of the steel breadboard.

The discontinuity may possibly be explained by an expansion of the steel and aluminum stage, due to a 1 K rise in stage temperature, probably caused by prolonged thermal contact of the operator with the stage armature during the rewinding procedure. A 1 K temperature rise would induce a 1.4 μm expansion between the retro-reflector of interferometer X and the quill of GLI CD, decreasing the summed lengths by this amount. For the same temperature change, the quill of GLI CA would be displaced by 0.4 μm along the line of sight. Since the stage temperature was not independently measured, such a hypothesis cannot be proved.

The discontinuity may alternatively have arisen from a crucial component

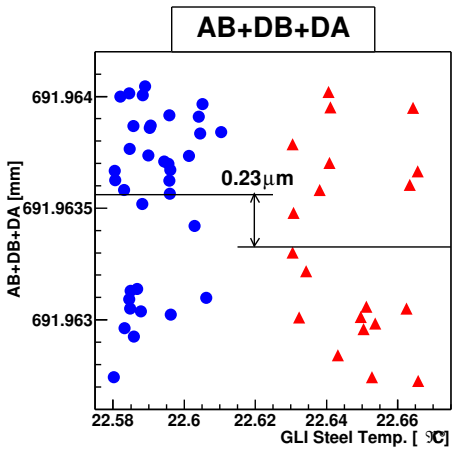

Figure 4.6: *The sum of lengths for GLIs forming the triangle ABD versus the scan averaged steel optical breadboard temperature. The blue circles and red triangles represent data before and after the stage was rewound. No significant discontinuity was observed.*

on the stage being touched during the rewinding procedure. It is highly unlikely that the discontinuity arose from systematic change in the prototype FSI system, because all measured lengths would have been affected, which was not the case. No significant discontinuity was present in the lengths of GLIs not convergent on node C, as shown by a plot against temperature of the sum of lengths of GLIs between nodes A, B and D, in Figure 4.6.

Due to the presence of the discontinuity, the last 20 FSI scans were discarded in the following analysis. For the first 35 FSI scans the maximum temperature change was only $\Delta T = 0.027$ K, so thermal effects were assumed negligible and were not studied further.

4.3.6 Calibration

To reliably reconstruct the grid geometry from real FSI measurements the SIMULGEO model needed to be calibrated by defining the offsets of each quill and retro-reflector from the grid node of the jewel. Only offsets along the GLI direction were considered, since any offset of the measurement node in the direction perpendicular to the GLI, would have only a very small effect on the calibration and reconstruction[1].

In the Monte Carlo model of Section 4.3.3, perfect nominal offsets were specified. For the prototype grid the real offsets could only be estimated within ± 1 mm from engineering drawings, machining tolerances and ruler measurements. The largest uncertainty was on the location of a quill within its supporting slotted-sphere, in which the quill was manually glued. The location of the apex of the retro-reflector was estimated to within ± 0.5 mm.

Jewel Offset Calibration Procedure

A special model for calibration was written in SIMULGEO in which all grid nodes were fixed at the corners of a square of side length 254 mm exactly. All grid node degrees of freedom were fixed. The position of all retro-reflectors within the jewel were fixed at an offset estimated from engineering drawings. All 6 quills were free to translate along the line from the grid node of the jewel supporting the quill, to the opposite retro-reflector. The 6 GLI measurements were passed to SIMULGEO and the 6 quill offsets were reconstructed, for each of the 35 FSI scans. The mean quill offsets were taken as the calibration constants of subsequent reconstruction model.

[1]For a 1 mm orthogonal displacement of the quill, the difference in GLI length would be only ~ 1 μm. This is far less than the uncertainty on the retro-reflector location in the GLI direction.

This general procedure should enable initial calibration of the ATLAS FSI grid. In the simple square geometry, the procedure is equivalent to defining each quill offset by a subtraction of the mean GLI length and retro-reflector offset from 254 mm for the short GLIs (or from $\sqrt{2} \times 254$ mm for the diagonal GLIs).

Notably, the procedure pinned the average shape of the grid as exactly 254 mm square. This baseline shape should therefore have been reconstructed by the calibrated model. What the real shape of the grid was at the micron level, does not matter, to first order. In the context of ATLAS, the relative positions of the grid nodes and the modules will not be known to microns: FSI measurements are only needed to precisely measure differences from a baseline shape of the SCT, determined by the X-ray, as discussed in Section 5.5. What does matter, is that the same shape is reconstructed if one or more FSI measurements fail; this is the main reason for the calibration. The precisions on the grid nodes should worsen if some GLI measurements fail, but the reconstructed shape should remain consistent if the reconstruction model has been correctly calibrated. The importance of correct calibration is demonstrated in Section 4.3.8.

Calibration of the model FSI measurement precision

The precision on each FSI measurement needed to be defined in the SIMULGEO model to reconstruct the grid. Naïvely all measurement precisions could have been defined as 1 μm. However, if the actual FSI precisions differed[2], two problems would arise:

- the precisions predicted by the error propagation would fail to match the reconstructed precisions.
- the incorrect ratio of precisions would be relied upon in the adjustment procedure of SIMULGEO resulting in a poor estimate of the reconstructed grid node co-ordinates.

The measurement precisions were therefore estimated for inclusion in the reconstruction model as follows, for two types of GLI measurements:

For the 3 GLIs not convergent on node C, the true GLI lengths were assumed constant for all 35 FSI scans. For each GLI length measurement, x_i, the precision was taken as the unbiased estimate of the population standard

[2]The ratio of measurement precisions affects the reconstruction.

deviation,

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2} \tag{4.2}$$

where

$$\overline{x} = \sum_{i=1}^{N} x_i/N \tag{4.3}$$

for N=35 FSI scans.

For interferometer X and the 3 GLIs convergent on node C, the true lengths were not constant due to the stage translation between groups of 5 FSI scans. The measurement precision was therefore estimated by,

$$s = \sqrt{\frac{1}{M}\sum_{j=1}^{M}\frac{1}{N-1}\sum_{i=1}^{N}(x_{ij} - \overline{x})^2} \tag{4.4}$$

for M=7 groups of N=5 FSI scans. The measurement precision was assumed independent of the measured length because the stage translations were relatively small.

The measurements precisions estimated by the above methods are presented in Table 4.3. The precisions differ primarily due to the amount of W-mode present in the return signal from each quill and also due to the retro-reflector quality. Importantly, all FSI measurement precisions were far smaller than the ATLAS requirement of ± 1 μm.

GLI	Mean length [mm]	Estimated σ measurement precision [nm]
AB	195.138457	$\pm$240
BC	197.393268	$\pm$220
CD	195.365242	$\pm$295
DA	195.926139	$\pm$125
CA	301.495481	$\pm$259
DB	300.898964	$\pm$257
X	118.818589	$\pm$202

Table 4.3: *GLI measurement precisions for the square grid, estimated using Equations 4.2 and 4.4.*

4.3.7 Reconstruction

The grid node co-ordinates were reconstructed in SIMULGEO with the calibrated model and are plotted in Figure 4.7 for each FSI scan. As for the Monte Carlo study, node A had no scatter since it defined the origin and node B had no scatter in the Y co-ordinate because it defined the X-axis. The scatter of the X co-ordinate of node B is shown for clarity as a histogram in Figure 4.8.

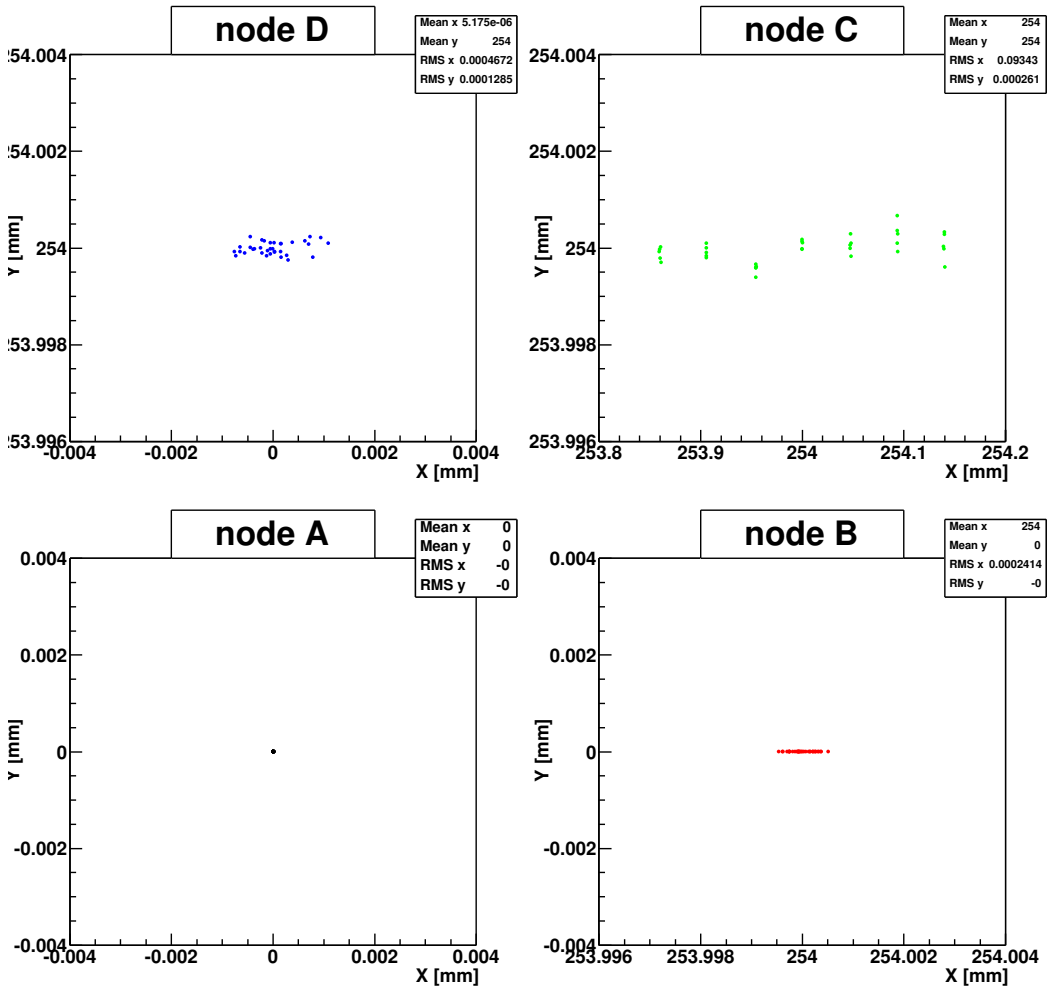

Figure 4.7: *The scatter in reconstructed node co-ordinates of the square grid for 35 FSI scans. The ranges of the axes have been retained from Figure 4.3 to allow a direct comparison with the scatter due to the required FSI precision for ATLAS. The extended X-axis of node C allows the 7 positions of the stage to be seen.*

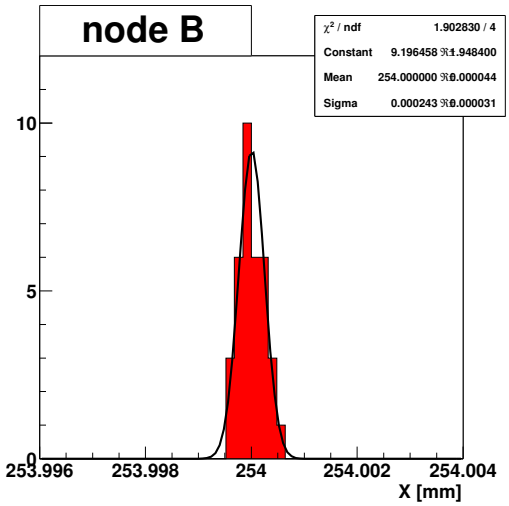

Figure 4.8: *The scatter in the reconstructed X co-ordinate of node B of the square grid. The χ^2/ndf for the fit was very sensitive to the histogram binning. A bin width was selected that gave fit values approximately in the centre of the range of fit values produced when a range of bin widths were tried. In the selected case the χ^2/ndf was less than 3.36/4 (the 50 % point in the χ^2 distribution for 4 degrees of freedom).*

As expected, the mean values of the reconstructed node co-ordinates matched those set during the calibration. Importantly, the scatters on the reconstructed node co-ordinates were found to be well within the ATLAS requirements, represented by the scatters of the Monte Carlo test in Figure 4.3, in which a 1 μm measurement precision was assumed. This was due to the sub-micron measurement precisions of the real FSI measurements.

A further error propagation was performed with the calibrated model, which included the measurement precisions from Table 4.3, to refine the predicted precisions. The reconstructed precisions were estimated from the 35 sets of reconstructed co-ordinates. The reconstructed precisions were calculated with Equation 4.2 for nodes B and D, which were assumed stationary and with Equation 4.4 for node C, due to the stage displacements.

The refined predicted precisions were compared with the reconstructed precisions, as in Figure 4.9, and were found to generally be in close agreement. Any disagreement, particularly apparent for the X co-ordinate of node D, arose mainly from the small number of grid measurements. The low statistic sample produced non-Gaussian distributions of the measured lengths. This induced an error in the estimate of the measurement precisions, that propagated though the grid in the error propagation to adversely affect the predicted precisions.

The same non-Gaussian distributions of GLI lengths, also lead to non-Gaussian distributions of the reconstructed co-ordinates and hence induce errors on the reconstructed precisions. If more grid measurements had been performed, these errors should have reduced, as observed for the simulated data of Section 4.3.3.

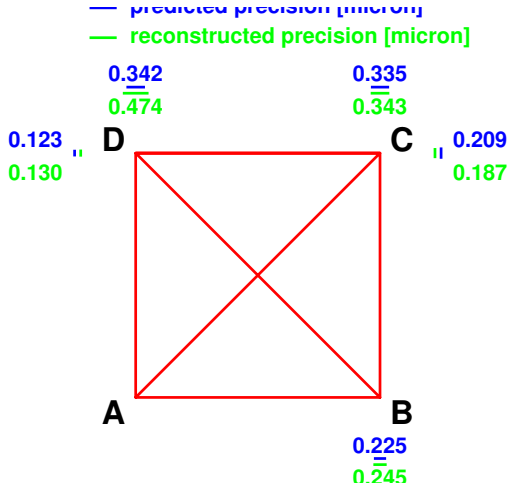

Figure 4.9: *Comparison of the predicted and reconstructed precision on the reconstructed grid node co-ordinates. The prediction was based on measurement precisions estimated from the measured lengths of each GLI.*

4.3.8 Grid Self-consistency

The redundancy in the grid was used to check the reconstruction process by systematically removing one of the 6 GLI measurements from the analysis. The 5 DoF in the model were therefore solved with only 5 measurements, so the grid was just constrained. The solution for one set of FSI measurements was unique and independent of the ratio of measurement precisions.

Plots of the reconstructed co-ordinates of nodes B, C and D are shown in Figures 4.10, 4.11 and 4.12 respectively, in which the lower plots correspond to a different combination of 5 GLI measurements. It was found that the nodes were still successfully reconstructed to within the requirements for ATLAS despite the removal of any one of the 6 GLI measurements.

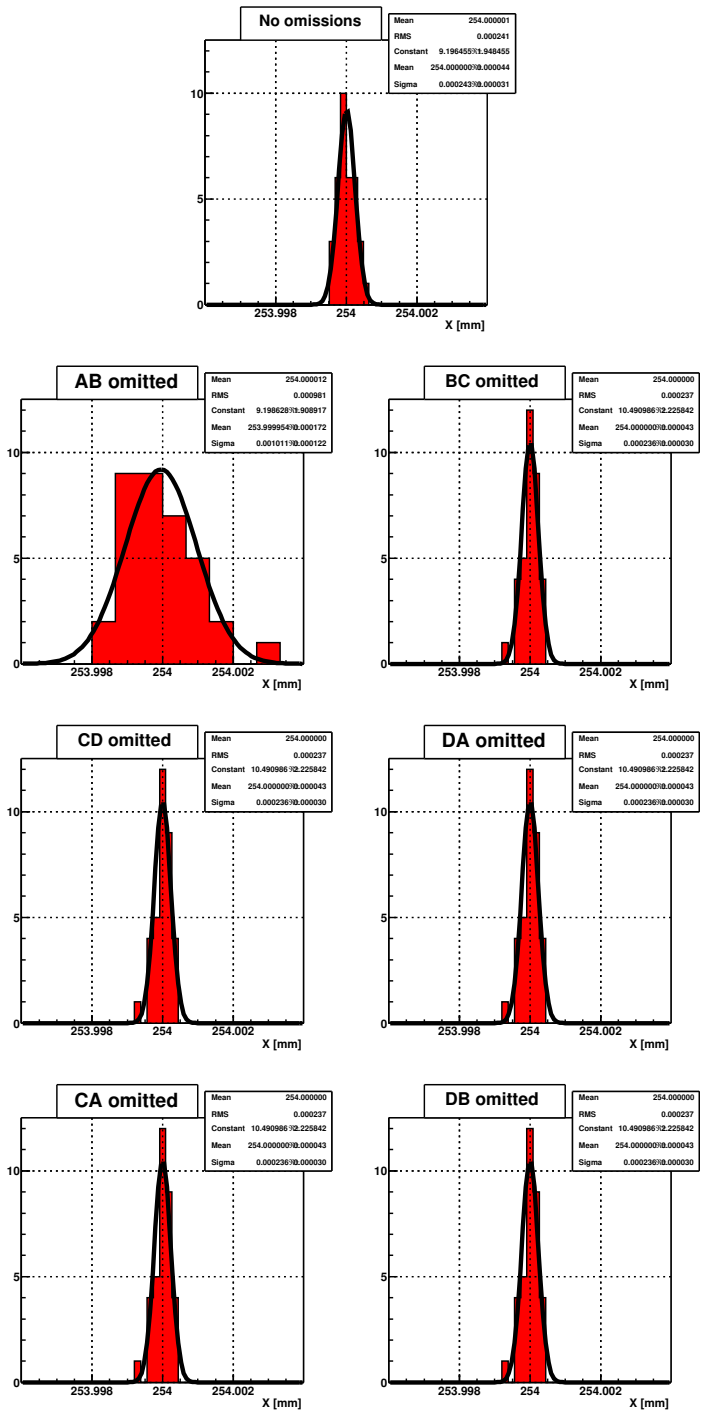

Figure 4.10: *The X co-ordinate of node B of the square grid, reconstructed with models having the combination of GLIs that omitted the GLI stated in the plot title. Since B defined the X-axis in the model, the Y co-ordinate had no scatter.*

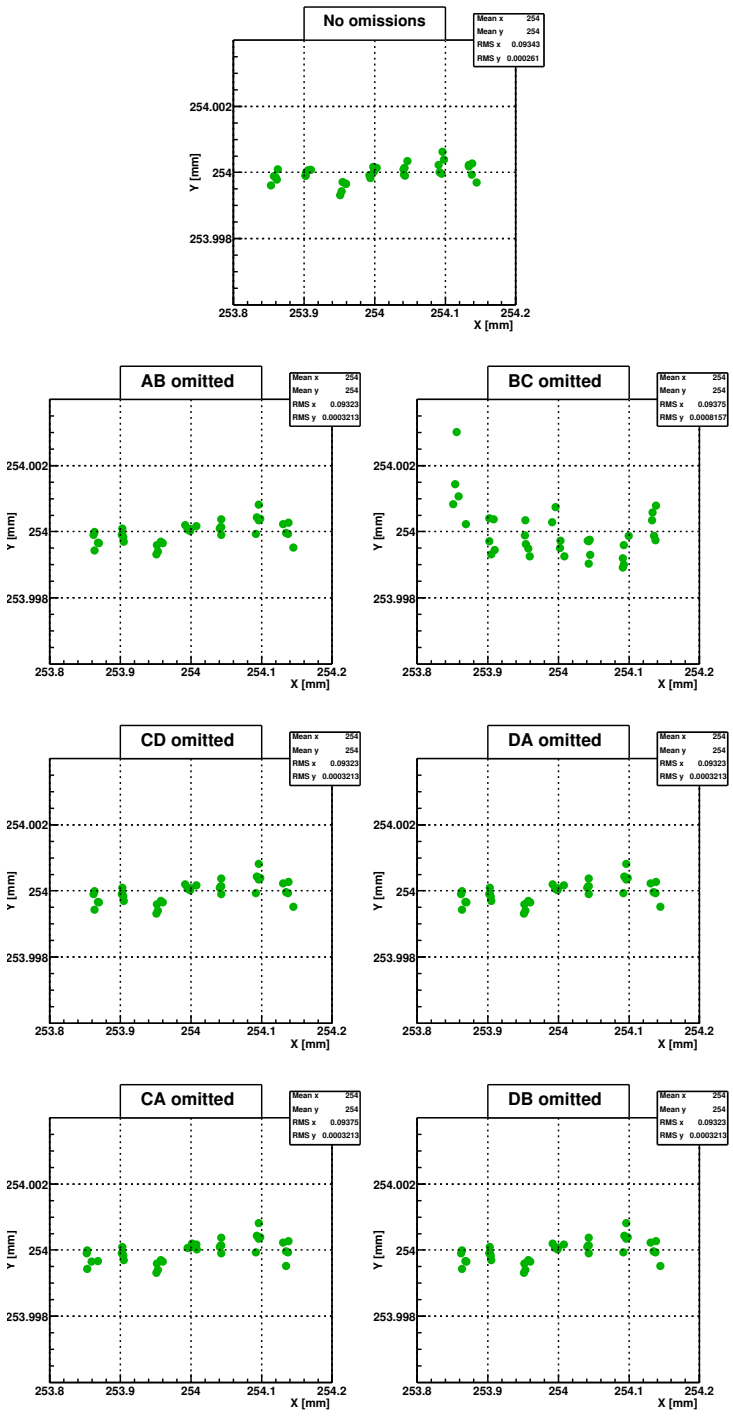

Figure 4.11: *The reconstructed co-ordinates of node C of the square grid, showing the 7 positions of the motion stage on which the node was mounted. The GLI stated in the plot title was omitted in the analysis.*

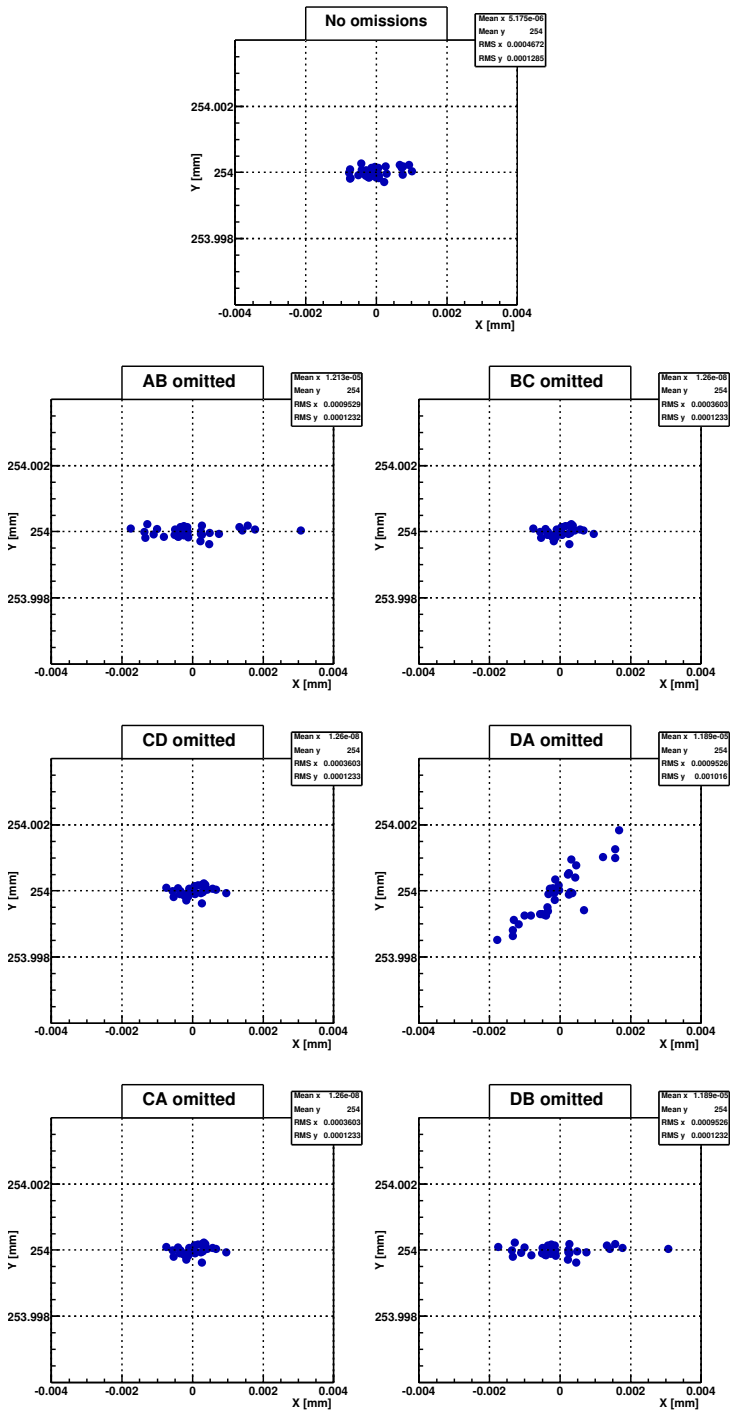

Figure 4.12: *The co-ordinates of node D of the square grid, reconstructed with models having the combination of GLIs that omitted the GLI stated in the plot title.*

Crucially, the mean reconstructed node co-ordinates for all analyses that omitted a GLI, successfully matched those reconstructed when all 6 GLIs were included. This was due to an appropriate calibration of the jewel offsets.

Miscalibration

The importance of the calibration was demonstrated by a similar set of analyses in which the offsets of the FSI measurement nodes within the jewel were set to crude (± 1 mm) estimates from engineering drawings, manufacturing tolerances and ruler measurements of the components. The reconstructed node co-ordinates from this poorly calibrated model are shown in Figure 4.13, again including the analyses that omitted one GLI. The widths and correlations of the local scatter in reconstructed node co-ordinates were similar to those for the calibrated model for each combination of GLIs. However, the mean values of the reconstructed node co-ordinates vary discontinuously on the order of millimetres (comparable to the miscalibration) as each GLI is systematically removed from the analysis.

These observations can be explained by the following analogy. Consider the GLIs as springs connecting the measurement nodes of the jewels. The spring constants are derived from the GLI measurement precisions. During a reconstruction, the grid nodes move until the forces between the springs balance, determining the shape of the grid.

In this analogy, during an appropriate calibration the jewels are pinned in a given configuration, while the springs are allowed to relax to their natural length (the mean GLI length) and the offsets of the measurement nodes from the jewel centres (the grid nodes) are recorded. The ends of the spring are then fastened into the jewel at the calibrated offset positions. When the grid is subsequently reconstructed from the mean GLI lengths, the grid settles to the shape in which it was pinned, corresponding to the natural length of the springs. If one GLI was removed from the reconstruction, the grid would still converge on the calibrated shape. Small perturbations of the spring lengths would result in corresponding changes to the reconstructed grid shape, allowing shape difference to be measured.

In the analogy, if poorly calibrated offsets were used to fasten the springs into the jewel, the shape reconstructed from the mean GLI lengths would differ from that of the calibrated model and internal tensions and compressions of the springs would occur. Small perturbations in the GLI lengths would however cause corresponding changes to the reconstructed shape. To first order, the poorly calibrated model fulfils the ATLAS requirement of correctly reconstructing shape differences from all GLI measurements. However, if any GLI was removed from the analysis, the sudden imbalance in spring forces would

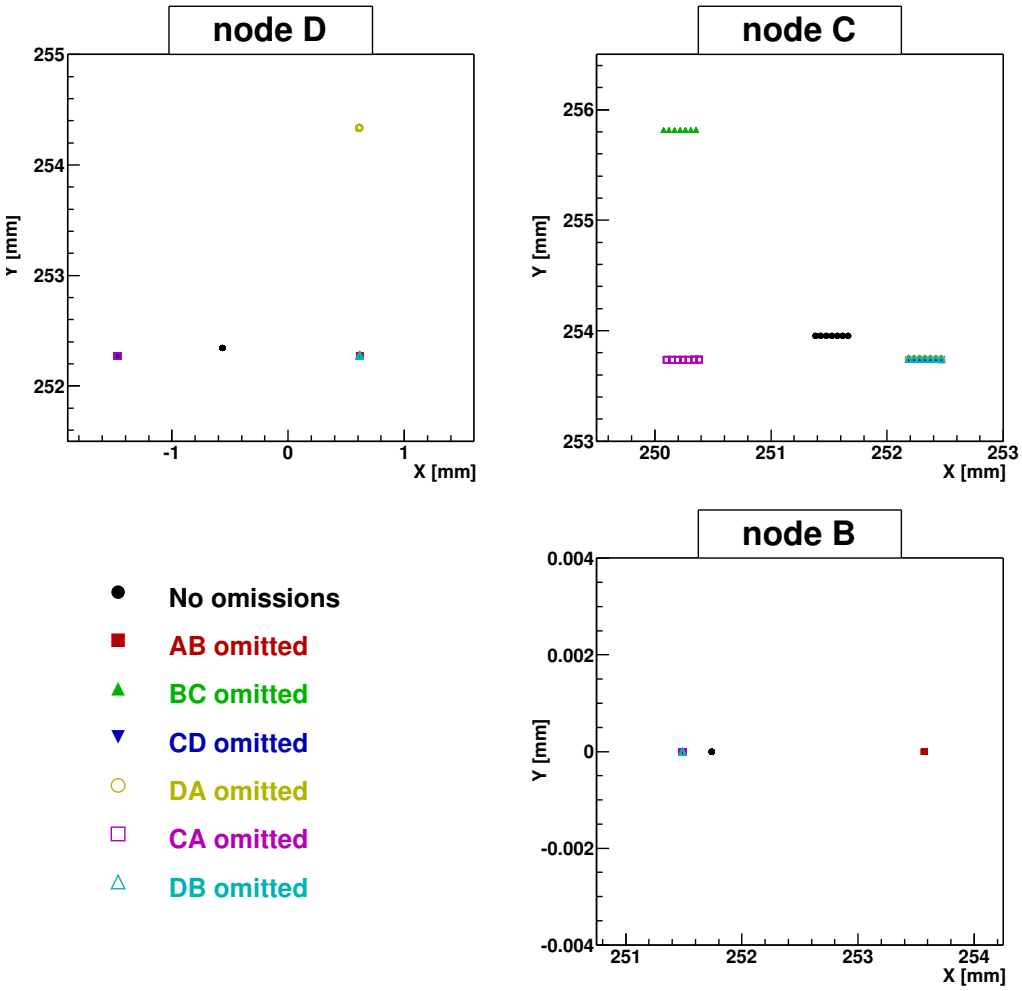

Figure 4.13: *The square grid node co-ordinates for the 35 FSI scans as reconstructed with a poorly calibrated model. As each GLI was systematically removed from the analysis the local reconstructed scatter remained small, but the mean values of the node co-ordinates varied discontinuously, on a similar scale to the miscalibration (±mm).*

drive the grid to an alternate shape, until equilibrium is again found. The difference in shape would be on the scale of the miscalibration. The shape of the grid would therefore discontinuously vary as each GLI is systematically removed, as was observed in the above analysis.

A poor calibration would therefore have serious consequences for the ATLAS SCT alignment system: the reconstructed shape would be severely altered if any FSI length measurement failed.

Error Propagation Comparison

For the appropriately calibrated model, the scatter on the reconstructed co-ordinates varied as each GLI was systematically removed from the analysis, as in Figures 4.10, 4.11 and 4.12. To check this variation in scatter, a set of refined error propagations were performed using the measurement precisions calculated in Section 4.3.6, for the different combinations of 5 GLIs. The reconstructed precisions were calculated for each combination of 5 GLIs by the method of Section 4.3.7. The refined predicted precisions are compared with the reconstructed precisions for each GLI combination in Figure 4.14.

The comparison revealed that for some node co-ordinates and GLI combinations, the predicted and reconstructed precisions agree to the nanometre; whereas for the other node co-ordinates, the predicted and reconstructed precisions differ by up to 450 nm. The following pattern was observed: only where more than one GLI was directly used to reconstruct a node do the predicted and reconstructed precisions not exactly agree.

These observations are explained by the identical estimation methods for the GLI measurement precisions and the reconstructed measurement precisions, operating on the low statistic sample of 35 grid measurements. As explained in Section 4.3.7, the low statistic non-Gaussian distributions of the measured GLI lengths induce errors in the estimates of the measurement precisions and hence the predicted precisions. Also the non-Gaussian GLI length distributions create non-Gaussian distributions in the reconstructed co-ordinates and hence the reconstructed precisions. In the special case that only one GLI length is used to directly reconstruct a node co-ordinate, then the GLI length and node co-ordinate distributions must match exactly. The predicted precisions and the reconstructed precisions therefore also match exactly, because they are calculated by the same method from identical distributions.

For example, from Figure 4.14 the predicted and reconstructed precisions on the X co-ordinate of node B are both 0.240 μm[3] for all combinations in which only GLI AB directly determines the parameter. The precisions differ only when AB is omitted and the X co-ordinate is determined via the other GLIs.

When all GLIs are included in the reconstruction, no exact matches are found between the predicted and reconstructed precisions of any node co-ordinate. With all GLIs, the grid is over-constrained so the free parameters in the model converge on the best fit in the adjustment procedure. No node co-ordinates therefore have the same distribution as any GLI length, so the predicted and reconstructed precisions should differ for the low statistics sample, as was observed.

[3]The estimated measurement precision on GLI AB is also 0.240 μm, as in Table 4.3.

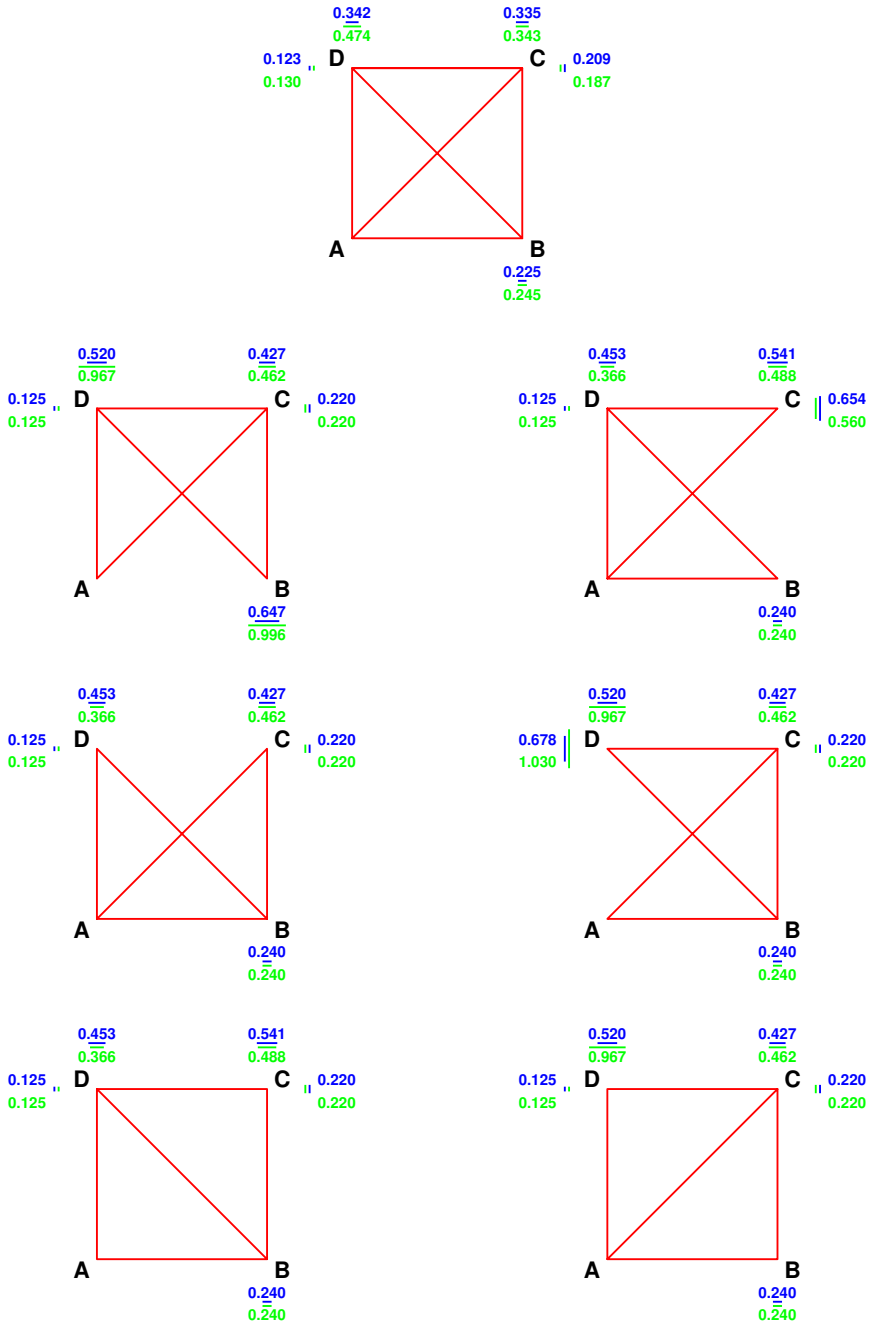

Figure 4.14: *Comparison of predicted and reconstructed precisions on the 5 free parameters of the square grid, for each model that omitted the GLI as shown.*

A few reconstructed precisions are larger for the analysis including all GLIs than for some combinations of 5 GLIs. However, the co-ordinates for the over-constrained model provide the best estimate of the shape differences of the grid.

4.3.9 Linearity Test

The linearity of individual FSI measurements has been verified [Coe01]. The linearity in reconstructing the position of grid nodes from combined FSI measurements was checked for the square grid, as described in this section.

The linearity on reconstructing node C was checked as the stage that supported node C was translated. The precision on reconstructing node C from the FSI grid measurements was found to be better than the resolution of around ± 1 μm on reading and setting the stage. The reconstructed X co-ordinate of node C was therefore compared, as in Figure 4.15, with the seventh FSI interferometer labelled 'X' that independently measured the stage translation to ± 0.2 μm (from Table 4.3). The 7 stage positions can be seen, corresponding to 6 full turns of the differential translation drive. A straight line fit to the data is shown to emphasize the important absence of any observable non-linear trend. The residuals from the best fit straight line have a standard deviation of $(0.364 \pm 0.044)\mu$m, as plotted in Figure 4.16. This spread is not significantly larger than the reconstruction precision on the node C X co-ordinate of 0.343 μm, estimated in Section 4.3.7. Therefore, any non-linearity due to the FSI grid measurement and reconstruction method was found to be too small to be resolved, over the 0.28mm translation range studied.

In Figure 4.15 the best fit straight line had an intercept compatible with 254 mm, expected from the calibration. The gradient was smaller than unity by 0.0024, probably due to:

1. non-collinearity of the X-axis of the reconstruction model, the stage translation and the measurement direction of interferometer X. The X-axis of the reconstruction model is defined by the line AB. A 0.5 mm offset of node A or B with respect to C in the plane perpendicular to the stage translation would be sufficient to cause the 0.7 μm observed difference between the reconstructed and independently measured X co-ordinate over the 0.28 mm range of the translation.

2. a differential thermal expansion of the optical breadboard and the plate on which the stage and interferometer X were mounted (although the plate was bolted to the optical breadboard).

Reason 1 is considered the most probable.

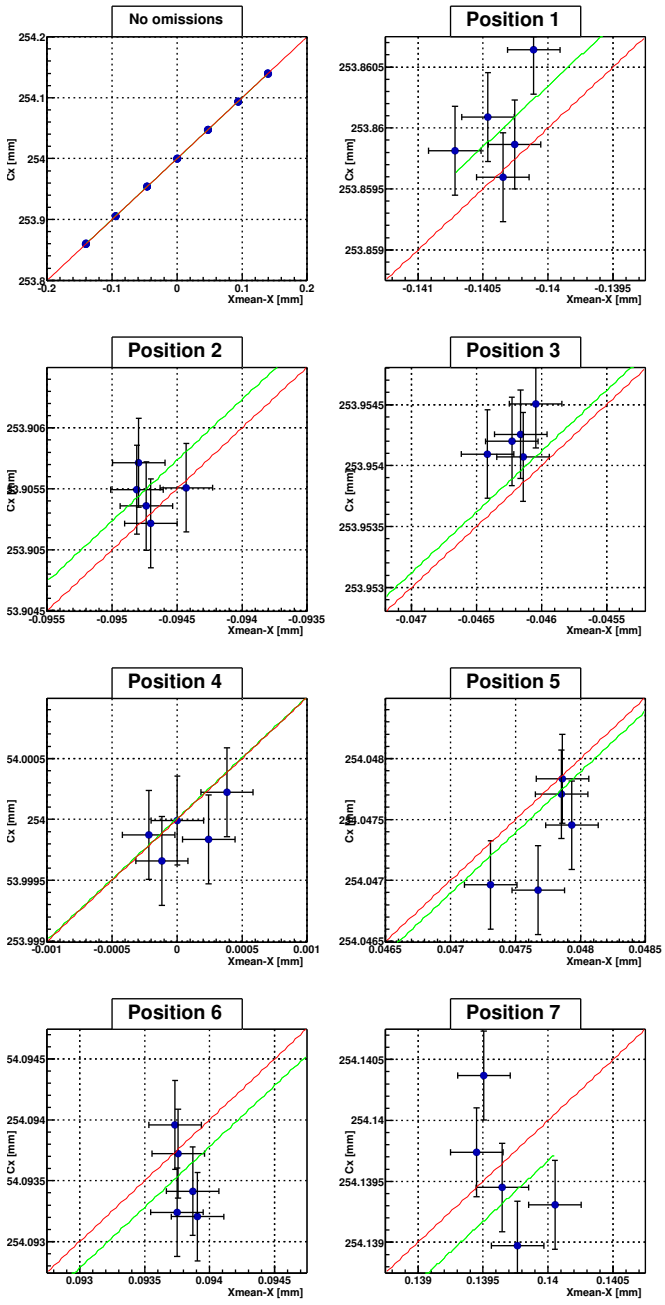

Figure 4.15: *The reconstructed X co-ordinate of node C for the square grid versus the independently measured stage position, where all GLIs was used to reconstructed the grid. Magnified regions at the seven stage positions are plotted. The best fit straight line is shown in green (gradient 0.997578). A straight line with a gradient of 1 is shown in red for comparison.*

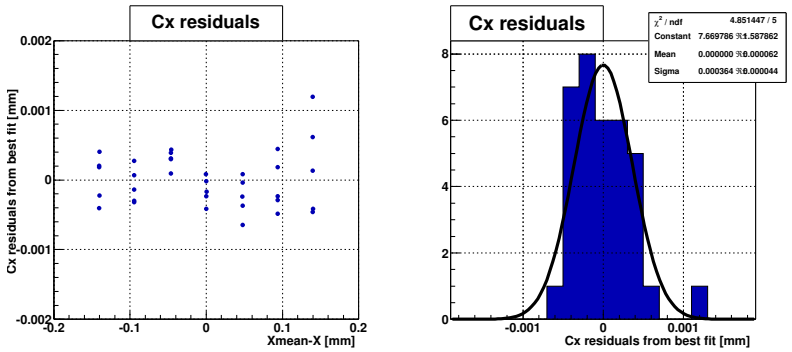

Figure 4.16: *The residuals of the reconstructed X co-ordinate of node C for the square grid from the best fit straight line shown in the plot of Figure 4.15.*

Linearity with omission of GLIs

The linearity was also assessed for the reconstruction models that omitted each GLI from the analysis. The reconstructed X co-ordinate of node C is plotted against the independently measured stage translation in Figure 4.17. The residual difference between the reconstructed X co-ordinate of node C and the independent measurement are plotted for each omitted GLI model in Figure 4.18. This is different to the plot of the Cx residuals from each best line fit; rather the X co-ordinate of node C is compared with a common baseline for all the GLI omission models. Notably, the reconstruction remains linear for all GLI combinations and the spread is given by the value labelled 'RMS y'. This spread includes the ± 0.2 μm error on the independent measurement.

Figure 4.19 shows the reconstructed Y co-ordinate of node C versus the deviation of the reconstructed X co-ordinate from the independently measured stage position, for each omitted GLI model. The Y-axes of Figure 4.18 form the X-axes of Figure 4.19. Figure 4.19 indicates the variation in scatter and correlations on the reconstructed co-ordinates of node C as each GLI is systematically removed from the analysis. The scatter in the X co-ordinate includes the error in the independent measurement (and any non-linearity). As for the other nodes, the scatter on node C was found to be well within the ATLAS requirements, despite any GLI being omitted from the reconstruction.

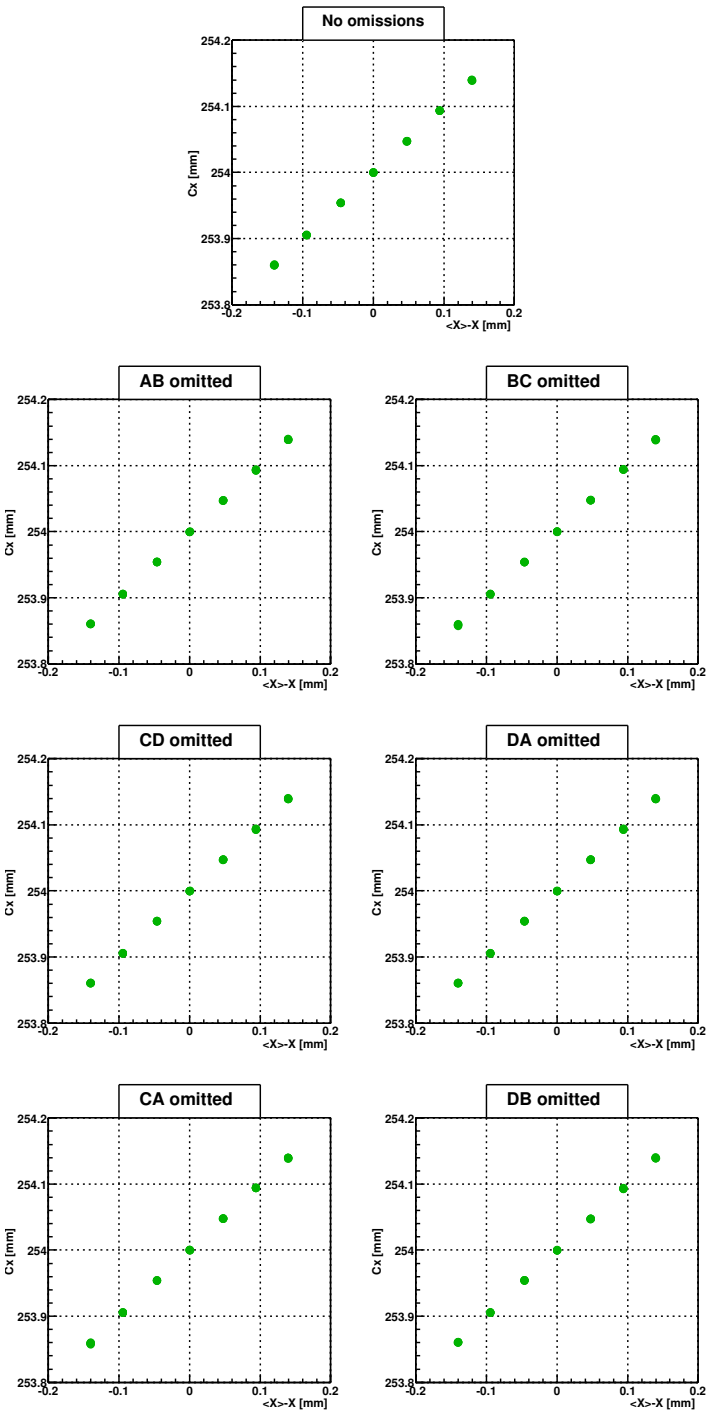

Figure 4.17: *The reconstructed X co-ordinate of node C of the square grid versus the independently measured stage position. The GLI stated in the plot title was omitted in the analysis.*

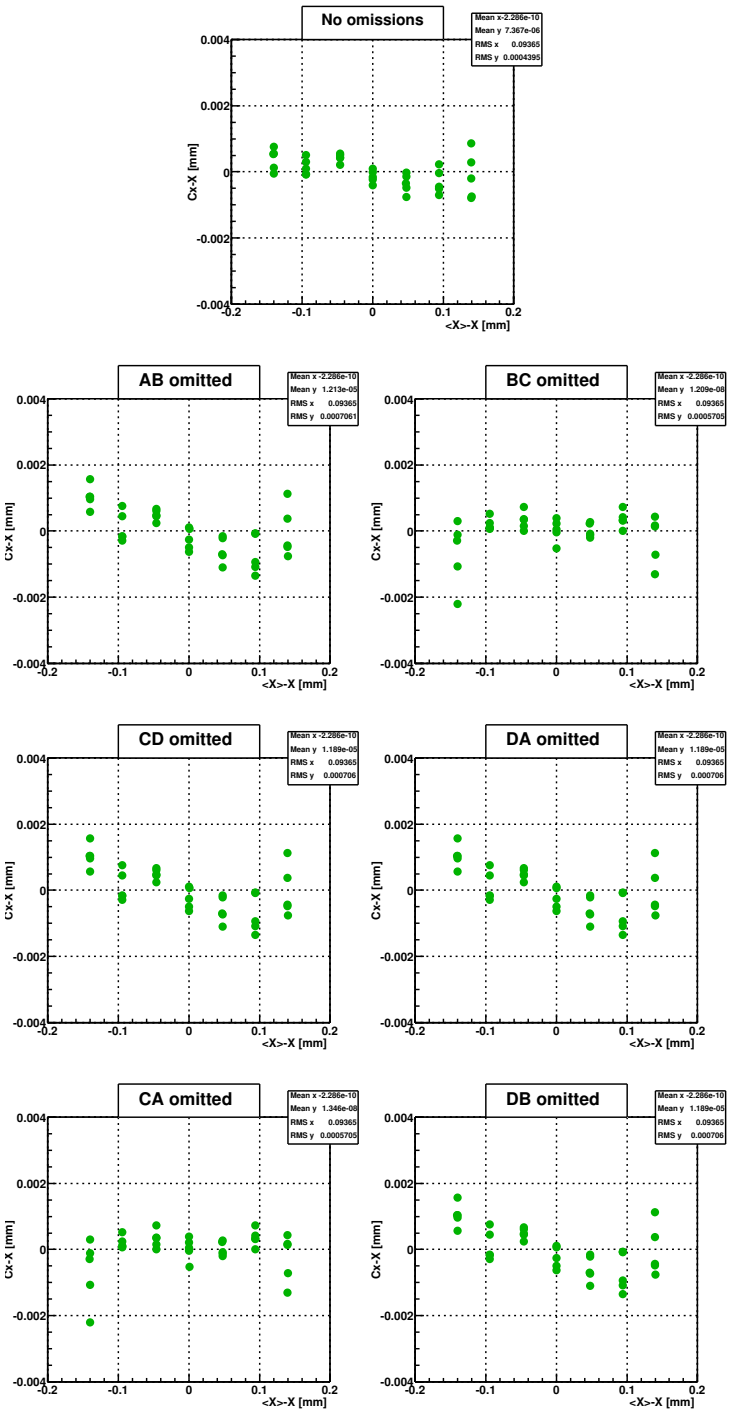

Figure 4.18: *The residual differences of the reconstructed X co-ordinate of node C of the square grid and the independently measured stage position.*

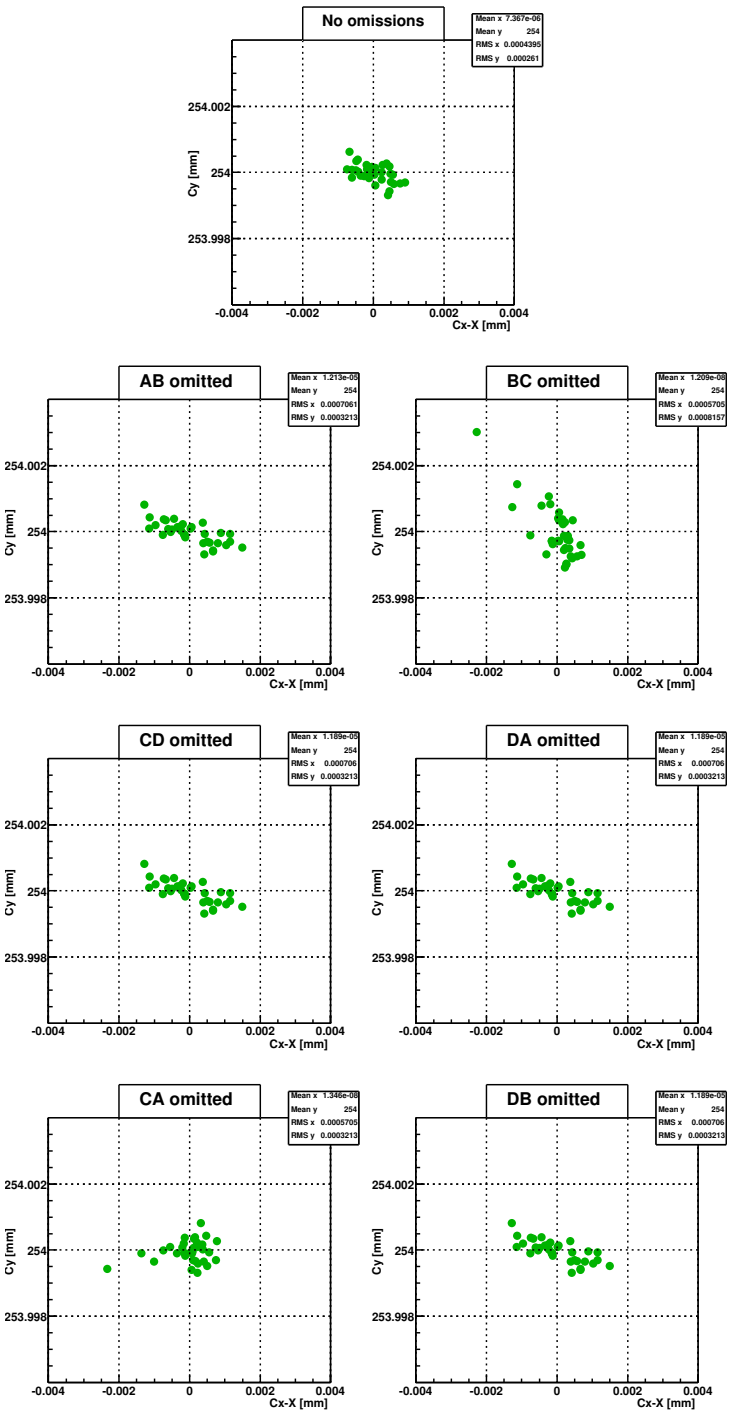

Figure 4.19: *The reconstructed Y co-ordinate of node C of the square grid versus the deviation of the reconstructed X co-ordinate from the independently measured stage position.*

4.3.10 Discussion and Conclusions

FSI geodetic grid shape measurement and reconstruction has been demonstrated for the first time with the square prototype grid. The precisions achieved on reconstructing the square grid node co-ordinates were on average a factor of 4.6 better than the precision required for ATLAS.

The presented data were reconstructed with a software model that had one particular DoF definition of the grid nodes. It could be argued that these DoF did not naturally suit the set-up: perhaps the DoF could have been defined in which the global co-ordinate frame coincided with the average position of the fixed nodes A, B and D. Such a model was written, along with several others to investigate the effect of different DoF definitions on the reconstruction. These results have not been presented here, but it was found that if the models had the same constraints and calibration, the DoF definitions did not significantly affect the reconstructed grid shape. The grid shapes reconstructed with any of these models for a set of GLI lengths, should be identical and related simply by co-ordinate transformations.

An appropriate calibration of the reconstruction model was crucial to consistently reconstruct the grid shape, when GLI measurements were individually omitted from the analysis. The miscalibration did not significantly affect the local scatter in reconstructed node co-ordinates, but the mean node co-ordinates varied on the order of the miscalibration. These results emphasize the importance of an appropriate calibration for the ATLAS FSI system, to ensure the reconstructed grid shape is not significantly altered if one or more GLI measurements fail.

The comparisons of the predictions and reconstructed precisions on the reconstructed grid node co-ordinates, for both the simulated and prototype grid measurements, demonstrated that the SIMULGEO models gave reliable predictions on FSI geodetic grid rigidity. This provided an important validation of SIMULGEO and its use for the simulations of the ATLAS FSI grids, detailed in Chapter 5.

The reconstruction with the square grid assumed that all four nodes were coplanar, so the grid was treated as two dimensional. The ATLAS FSI grids will be three dimensional and co-ordinate measurement of this type was investigated with the tetrahedral grid, as described in Section 4.4.

4.4 Tetrahedral Grid

4.4.1 Overview

The square grid was adapted to the tetrahedral geometry, described in Section 3.3.3, to allow investigations of grid node reconstruction in three dimensions. The grid was reconstructed with a common software model, described in Section 4.4.2, for all but one of the investigations. Initially the grid was measured repeatedly, without adjusting the grid shape, to allow the length measurement precisions to be determined and the reconstruction model to be calibrated, as described in Section 4.4.4. The node mounted on the motion stage was finely manipulated to test if the node position could be precisely reconstructed in three dimensions from FSI measurements.

4.4.2 Reconstruction Model

A SIMULGEO model of the tetrahedral grid was defined in which node A was the origin and the line from node A to node B formed the X-axis. The plane of nodes A, B and D defined the X-Y plane. These 3 nodes had the same 3 DoF as for the square grid model. Node C was free in all 3 translational DoF. The rotational DoF of the jewels were assumed to be fixed. The model thus had 6 DoF determined by 6 distance measurements, so the tetrahedral grid was just constrained.

The model was calibrated for reconstruction of real FSI measurements, as discussed in Section 4.4.4, although preliminary studies were performed with a model describing a perfect grid geometry.

4.4.3 Preliminary Reconstruction Studies

Error Propagations

An error propagation was performed with the SIMULGEO model, to predict the precision on the reconstructed node co-ordinates for length measurement precisions that exactly met the ATLAS requirement of 1 μm. The predicted precisions on the reconstructed node co-ordinates, shown in Figure 4.20, were found to be generally worse than the equivalent predictions for the square grid, shown in Figure 4.2. The difference is primarily due to the square grid being overconstrained and the tetrahedral grid, which had an additional DoF in Z on node C, being just constrained. The predicted precision on any node of the tetrahedral grid was therefore equal to, or worse than, the assumed measurement precision of 1 μm. The Z co-ordinate of node C was predicted to

be determined relatively poorly because of the small Z components of the GLIs convergent on node C, which resulted from the skewed tetrahedral geometry.

A second error propagation was performed with a model of the tetrahedral grid in which the Z co-ordinate of node C was rigidly fixed. Hence there were 5

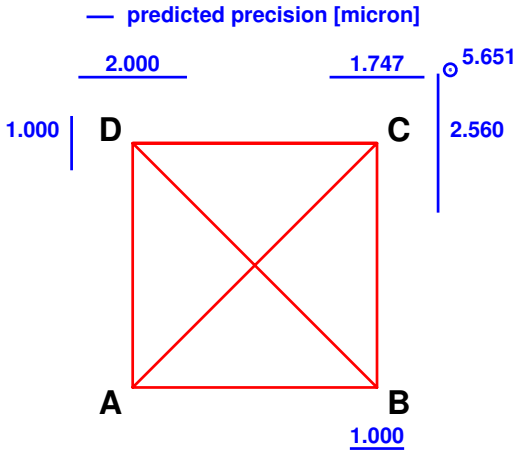

Figure 4.20: *An error propagation for the tetrahedral grid, assuming the 1 μm FSI measurement precision required for ATLAS. The predicted precisions on the 6 free parameters of model, including the Z co-ordinate of node C, are shown. In this model, the grid was just constrained.*

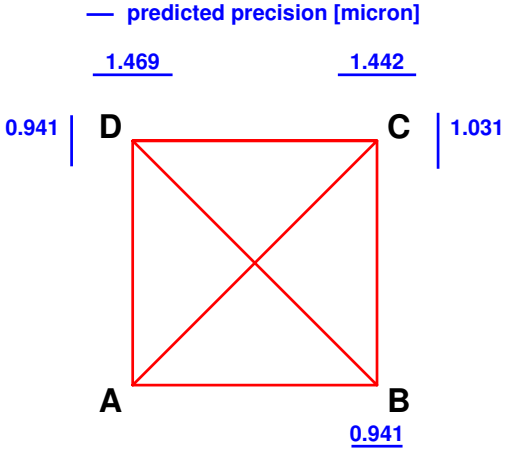

Figure 4.21: *An error propagation for the tetrahedral grid, assuming a 1 μm FSI precision and node C to be rigidly fixed in Z. The predicted precisions on the 5 parameters are shown next to the relevant node. In this model the grid was overconstrained.*

DoF that were overconstrained by the 6 GLI measurements. The redundancy improved the predicted precisions on the node co-ordinates, to those in Figure 4.21, so that some precisions were less than the assumed 1 μm measurement precision. The extra constraint assumed in this model would cause the incorrect X and Y co-ordinates of node C to be reconstructed, if node C moved in Z by more than a few microns.

Check with Monte Carlo

A set of simulated length measurements were reconstructed with the tetrahedral grid model, in which node C was free in all 3 translational DoF. The reconstruction process was checked and the correlations between reconstructed node co-ordinates were evaluated. The GLI lengths were calculated from an independent computer model that described a perfect geometry of the tetrahedral grid: the grid nodes were separated by 254 mm in X and in Y, and node C was raised above the plane ABD by exactly 100 mm. 10,000 copies of each GLI length were independently smeared with a Gaussian of width 1 μm to simulate length measurement errors. The grid nodes were reconstructed from each set of smeared lengths with SIMULGEO.

As expected, the reconstructed co-ordinates of node A had no scatter, because node A defined the origin, nor did the Y co-ordinate of node B, because AB defined the X-axis. The reconstructed co-ordinates of nodes B & D and of node C are plotted in Figures 4.22 and 4.23 respectively.

The mean positions of all reconstructed nodes matched those in the original model geometry. The reconstructed precisions were determined by a Gaussian fit to each node co-ordinate and are compared with the predicted precisions in Table 4.4. The precisions were found to be compatible, within the statistical fit

Grid node co-ordinate	Predicted precision [μm]	Simulated reconstructed precision [μm]	Fit error on reconstructed precision [μm]	Reconstructed minus predicted [μm]
Bx	1.000	1.004	$\pm$0.007	+0.004
Cx	1.747	1.741	$\pm$0.012	-0.006
Cy	2.560	2.539	$\pm$0.018	-0.021
Cz	5.651	5.590	$\pm$0.040	-0.061
Dx	2.000	2.000	$\pm$0.014	+0.000
Dy	1.000	1.004	$\pm$0.007	+0.004

Table 4.4: *Comparison of predicted and simulated reconstructed precisions on the tetrahedral grid node co-ordinates, assuming a 1 μm measurement precision required for the ATLAS FSI system.*

errors, estimated by Minuit [Jam75], demonstrating that the three dimensional reconstruction proceeded as expected.

The Monte Carlo check also revealed the correlations between certain free parameters. The correlations of the 3 co-ordinates of node C create an error ellipsoid, that has its major axis tilted from the vertical, approximately towards the centre of the grid, mainly as a result of the small Z components of the GLIs convergent on node C. The tilt is not directly towards the centre of the grid due to the asymmetric DoF definitions in the model.

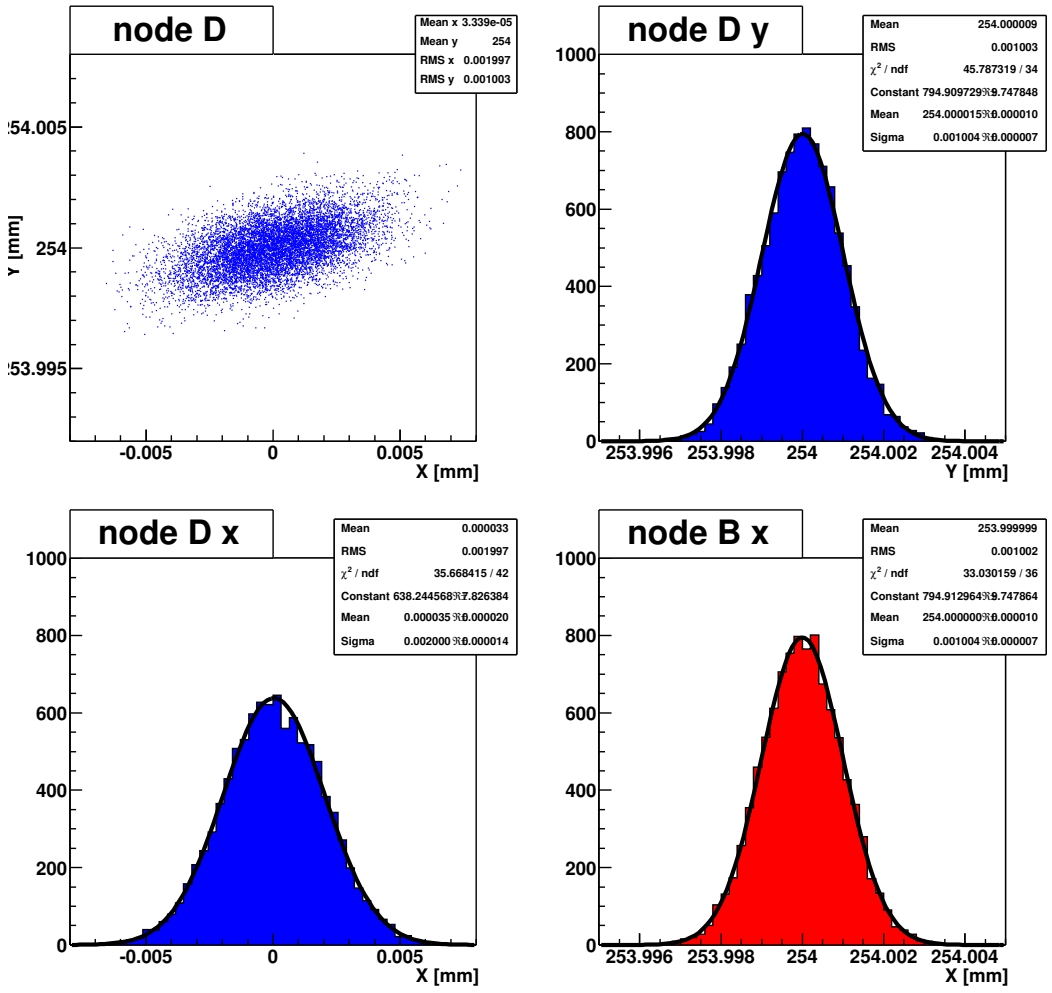

Figure 4.22: *The simulated scatter in reconstructed co-ordinates of nodes B and D of the tetrahedral grid, due to a 1 µm smearing of the GLI lengths, representing the required ATLAS FSI precision. Node B defined the X-axis in the model, so has no scatter in the Y co-ordinate.*

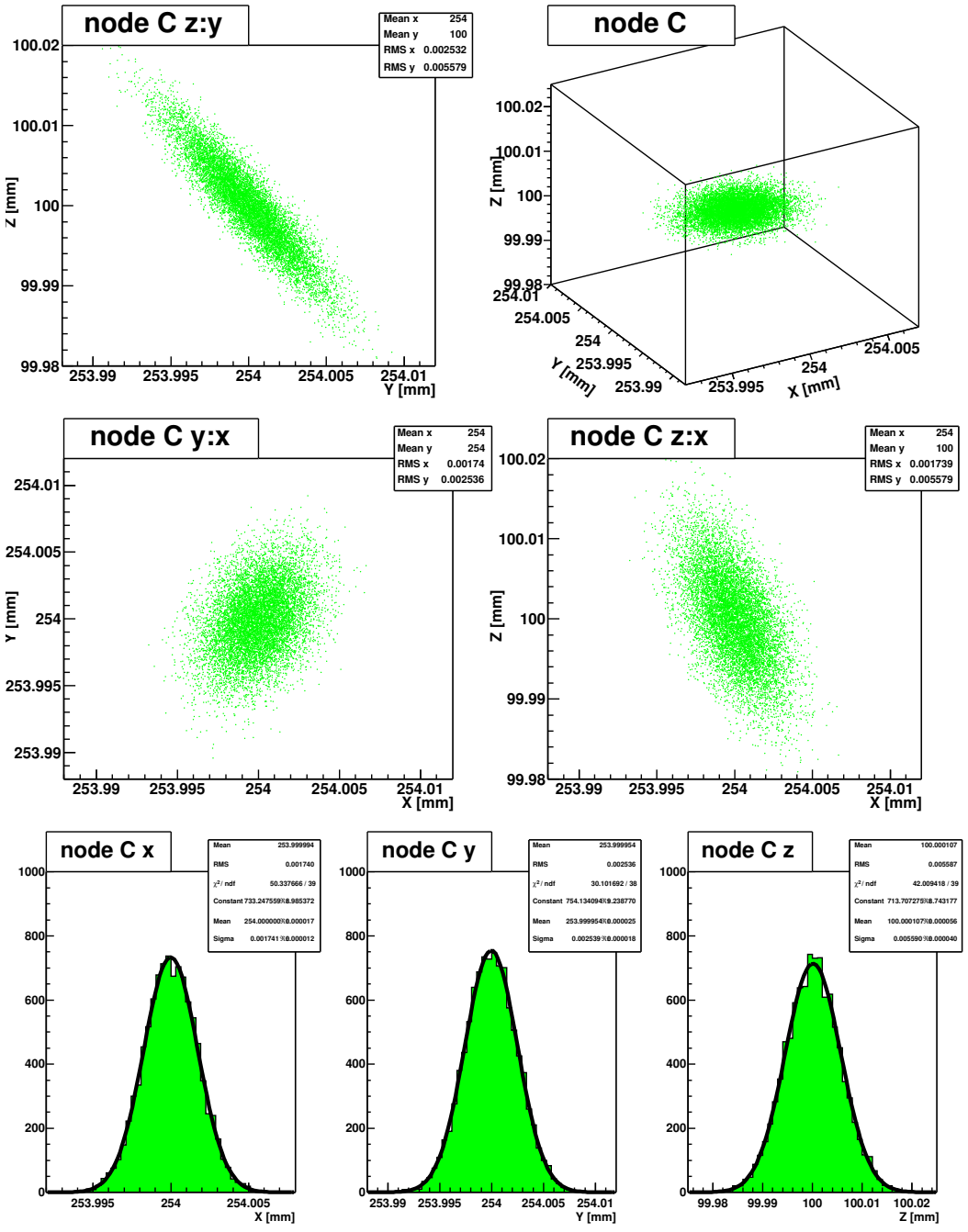

Figure 4.23: *The simulated scatter in the 3 reconstructed co-ordinates of node C of the tetrahedral grid, due to a 1 µm smearing of the GLI lengths. The correlations of the different node C co-ordinates are shown for clarity. The simulated reconstructed precisions were determined by a Gaussian fit to each node C co-ordinate.*

4.4.4 Calibration

Jewel Offset Calibration

As found for the square grid, to reliably reconstruct the grid geometry from real FSI measurements, the SIMULGEO model of the tetrahedral grid needed to be calibrated. The grid geometry had been adapted from the square configuration by elevating node C through 100.0mm, as described in Section 3.3.3. The quills in node C were reorientated and realigned with the corresponding retro-reflectors of the other 3 nodes. Due to geometry changes caused by the realignment, the offsets of each quill and retro-reflector from the grid node were not retained from calibration of the square grid, but were instead recalibrated as follows.

A special SIMULGEO model for calibration was written in which all grid nodes were fixed at the nominal positions. Nodes A, B and D were defined as for the square grid calibration, such that all retro-reflector locations were fixed and the quills were free to translate along the GLI directions. The 3 quills of node C were defined in distinct co-ordinate frames, as shown in Figure 4.24, that were secured in a fixed orientations to the frame of jewel C. The positions of the quills frames were set at a constant radius from the grid node, taken from engineering drawings, to simulate the fixed position of the quill-spheres. The orientations were defined by the declination angles of the GLIs diverging from node C,

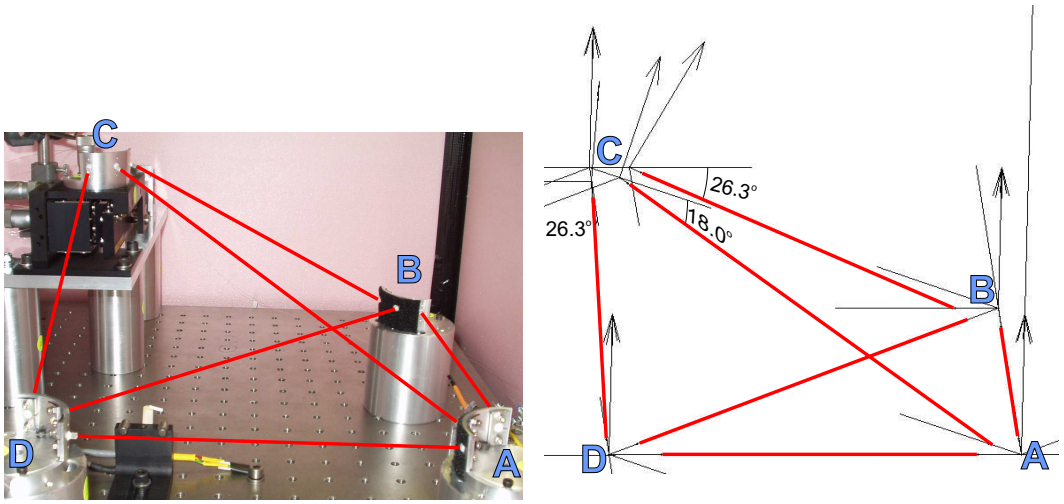

Figure 4.24: *The skewed tetrahedral geometry of the three dimensional prototype FSI grid and the equivalent rendering of the SIMULGEO model. The GLIs have been overlaid with red lines for clarity. The quills were free to translate along the GLI directions during the calibration.*

calculated (as ∼26.3° and ∼18.0° for GLIs BC or CD, and CA respectively) from engineering drawings and confirmed with ruler measurements. These 3 quills were released to translate along the line of the GLI, which was assumed to be from the centre of the quill-sphere to the apex of the opposite retro-reflector. This procedure therefore accounted for the large lateral offset on the quill measurement nodes, induced by the reorientation of the 3 quills of node C during realignment.

The model was calibrated with data from the 41 FSI scans of the stability test, described in Section 4.4.5. For each scan the 6 GLI lengths were passed to SIMULGEO and the quill offsets were reconstructed. The mean quill offsets were taken as the calibration constants.

Calibration of the model FSI Measurement Precision

The FSI measurement precisions required in the SIMULGEO model for grid reconstruction, were calculated for all GLIs with Equation 4.2, for N=41 FSI scans of the stability test. The estimated precisions are presented in Table 4.5. The precisions were found to differ from those previously estimated for the square grid, in Table 4.3. The differences are due to the effects of temperature, as discussed in Section 4.4.5, and the realignment of certain GLIs. The measurement precision of GLI AB was worse for this data set, as discussed in Section 4.4.5, although for all GLIs, the measurement precisions were far smaller than the ATLAS requirement of 1 μm.

GLI	Mean length [mm]	Estimated σ measurement precision [nm]
AB	195.138112	±554
BC	220.072195	±235
CD	217.962714	± 83
DA	195.926860	±181
CA	316.702073	±131
DB	300.898801	±220
X	118.900561	±433

Table 4.5: *GLI length measurement precisions for the tetrahedral grid, estimated from the stability test data using Equation 4.2.*

4.4.5 Stability Test

Measurement Procedure

An initial test was performed to assess the precision in reconstructing the tetrahedral grid while the geometry was stable. 41 consecutive FSI scans were recorded without adjustment of the stage supporting node C. The thermal insulation around the grid was not opened during these measurements. This data was used to calibrate the reconstruction models of all subsequent tests.

The measurements were recorded over 5.6 hours, during which time the temperature of the optical breadboard rose approximately linearly by 0.285K.

Reconstruction

The grid node co-ordinates were reconstructed for each FSI scan with the calibrated SIMULGEO model, in which node C was free to translate in the Z co-ordinate.

No scatter was found on the reconstructed co-ordinates of node A or on node B in Y, due to the model definitions previously described. The reconstructed co-ordinates of nodes B & D and of node C are plotted in Figures 4.25 and 4.26 respectively.

As expected, the mean values of the reconstructed co-ordinate matched those set during the calibration. The reconstructed precisions were estimated from the 41 sets of reconstructed node co-ordinates. All nodes were assumed to be stationary in this test, so the estimated reconstructed precisions were calculated with Equation 4.2 for all nodes.

Importantly, the precisions on the reconstructed co-ordinates were found to be on average a factor of 4.3 better than the precisions predicted when the required ATLAS FSI precision was assumed. The average improvement was less than the improvement of 4.6 for the square grid, because the reconstructed X co-ordinate of node B exhibited a particularly non-Gaussian distribution in this test. This was a direct result of the least precise GLI measurement, AB, as in Table 4.5.

A refined error propagation was performed in which the measurement precisions from Table 4.5 were assumed. The predicted precisions were compared with the estimated reconstructed precisions, as in Figure 4.27, and were found to generally be in close agreement. Any differences results from the non-Gaussian distribution of the low statistic sample of 41 FSI scans, as discussed in Section 4.3.7, and from any real relative node movements.

The correlated scatters in the reconstructed node co-ordinates, shown in Figures 4.25 and 4.26, approximately matched those found in the Monte Carlo

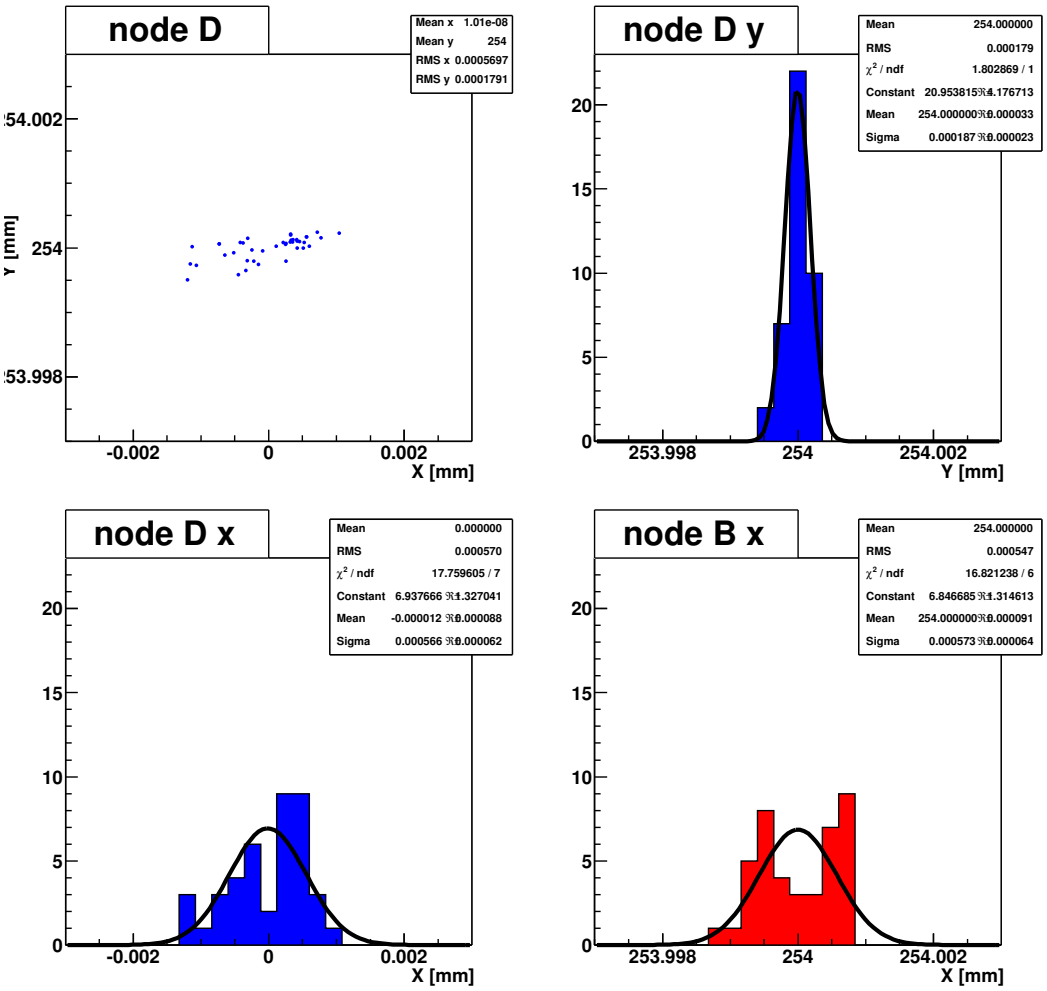

Figure 4.25: *The scatter in reconstructed co-ordinates of nodes B and D of the tetrahedral grid, for 41 FSI grid measurements of a stable grid. Node B defined the X-axis in the model, so has no scatter in the Y co-ordinate. The scatter on node B was found to be particularly non-Gaussian for this test.*

test shown in Figures 4.22 and 4.23, given the low statistics and the effects of real relative node movements, discussed below.

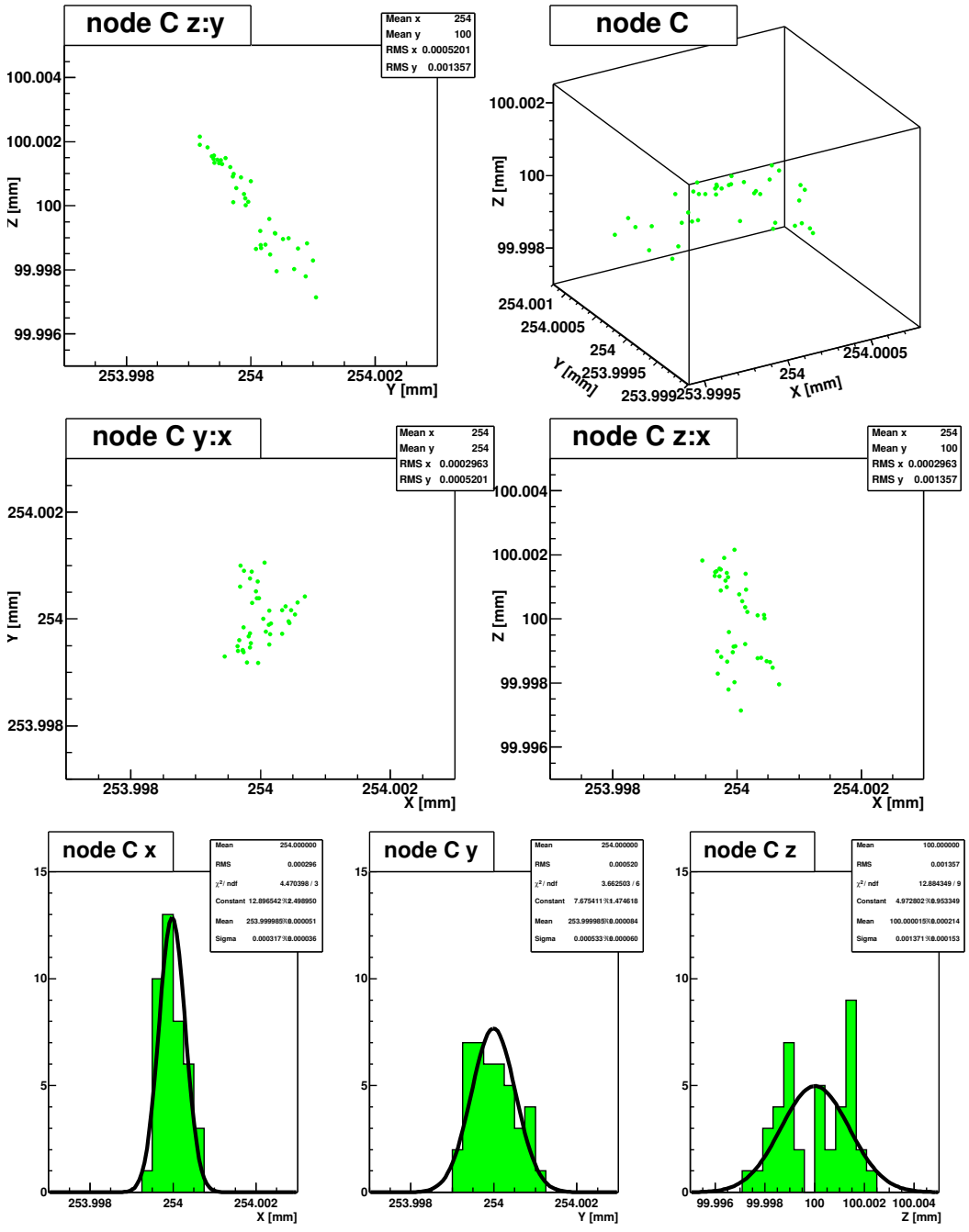

Figure 4.26: *The scatter in the 3 reconstructed co-ordinates of node C of the tetrahedral grid, for 41 FSI grid measurements, without adjusting the position of node C.*

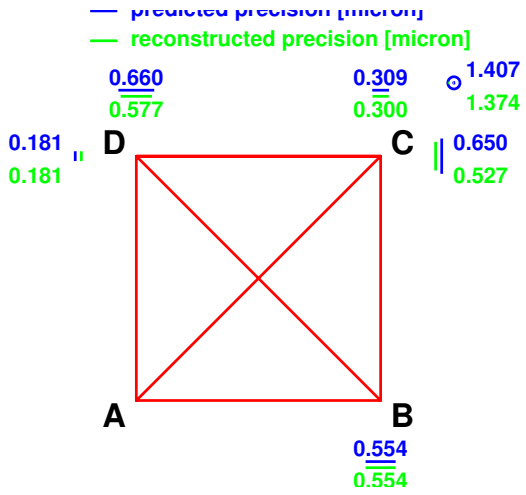

Figure 4.27: *Comparison of the predicted and reconstructed precision on the 6 reconstructed tetrahedral grid node co-ordinates. The prediction was based on measurement precisions estimated from the measured lengths of each GLI.*

Consideration of Thermal Expansion

The scatters on the reconstructed node co-ordinates also included any real relative movement of the nodes, that could for example have been caused by the thermal expansion of the set-up. Figure 4.28 shows the correlation of each reconstructed node co-ordinate with the temperature of the steel optical breadboard.

The increase in the Z co-ordinate of node C of over 3.5 μm is too large to be explained by a simple correlation with temperature. The differential expansion of the aluminium pillars on which node A and node C were mounted, would induce an increase in the Z co-ordinate of node C of:

$$\begin{aligned} \Delta L &\approx \alpha_{aluminium} L \Delta T \\ &\approx 0.66 \ \mu\text{m} \end{aligned} \tag{4.5}$$

where:
coefficient of expansion $\alpha_{aluminium} = 23.1 \times 10^{-6}$ K^{-1} [Kay89],
distance in Z between nodes A and C, $L = 100$ mm,
measured temperature change of optical breadboard, $\Delta T = 0.285$ K.

The reconstructed X and Y co-ordinates of node C, both decreased slightly with temperature, in contradiction to the increase of about 0.8 μm expected from the temperature rise, as calculated from Equation 4.1.

The increase in the reconstructed Y co-ordinate of node D was compatible with the 0.8 μm expected from thermal expansion over the distance from node

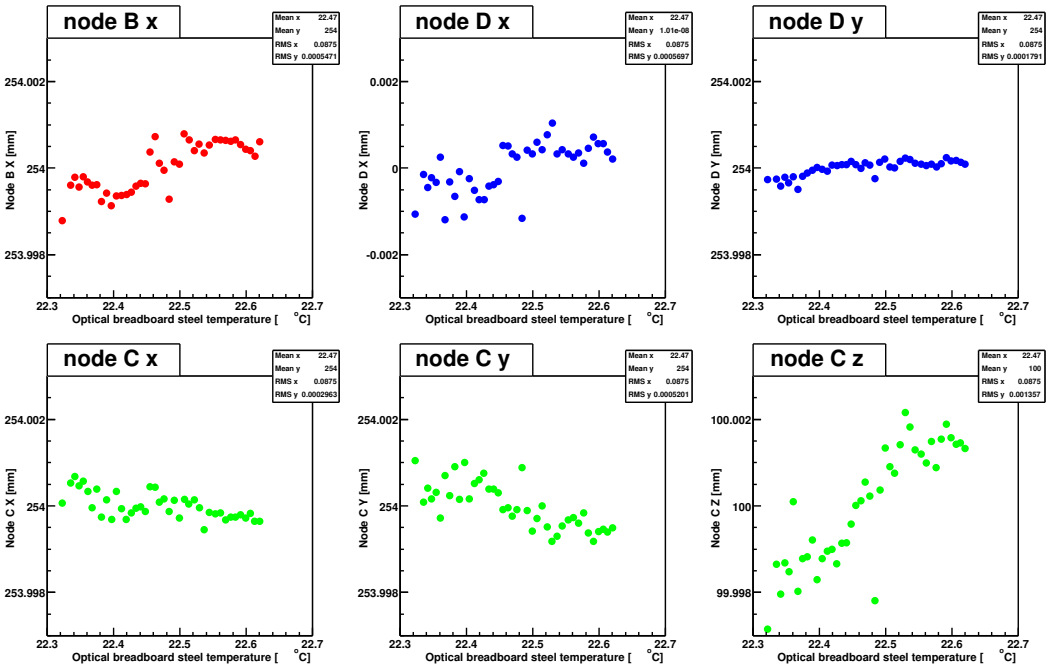

Figure 4.28: *Correlation of the 6 reconstructed node co-ordinates with the temperature of the steel optical breadboard. The co-ordinate axes of the plots share a common 6 μm range to aid comparison.*

A to D. The X co-ordinate of node D should have remained fixed with respect to node A, for a linear expansion of the breadboard, but instead increased slightly. Importantly, the reconstructed X co-ordinate of node B increased in a non-linear fashion by around 1.5 μm, which is more than the 0.8 μm expected from thermal expansion.

The observations may be explained solely by the poor measurement of GLI AB. The temperature correlations of the 6 measured GLI lengths are shown in Figure 4.29.

The reconstructed positions of nodes B and D with respect to node A were determined entirely from the GLIs within the triangle ABD, because the grid was just constrained. The X co-ordinate of node B was determined directly from the GLI AB, as apparent from the identical distributions of Bx and AB shown in the upper left plots of Figures 4.28 and 4.29 respectively. Also, the Y co-ordinate of node D is determined directly from the GLI DA, for small Y movements of node D, as shown by the identical distributions in the upper right plots of the same Figures. The X co-ordinate of node D was determined via DB from the reconstructed position of node B. Therefore the non-Gaussian distribution on AB was propagated to and included in the reconstructed X

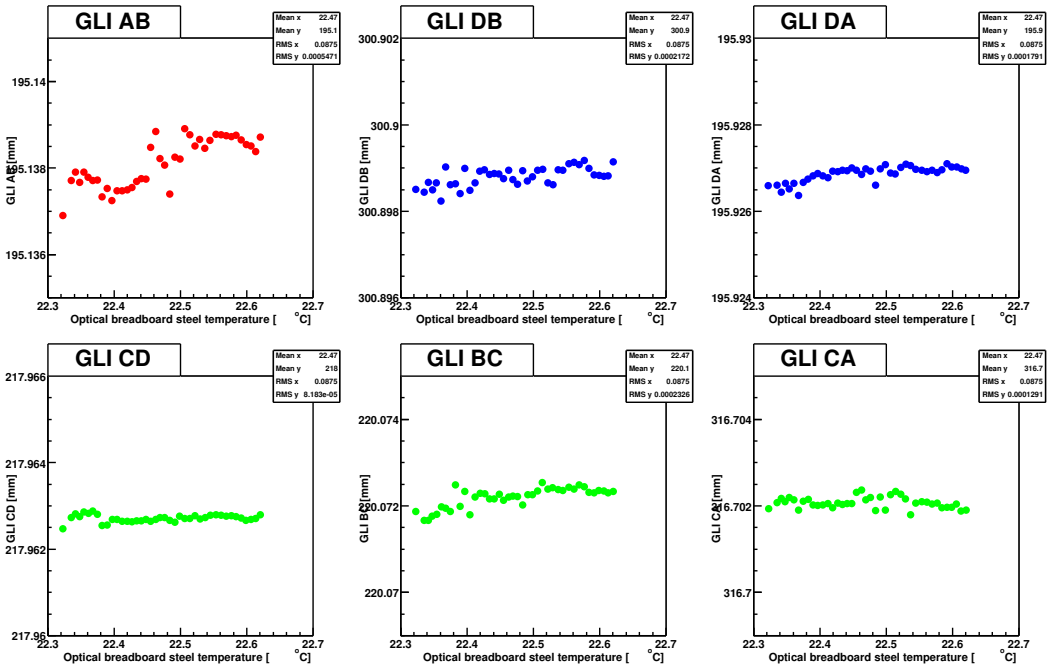

Figure 4.29: *Correlation of the 6 measured GLI lengths with the temperature of the steel optical breadboard. A scale of 6 µm has been retained from Figure 4.28 to allow a direct comparison.*

co-ordinate of node D.

The measured length of GLI CD varied very little with temperature. Therefore the erroneous non-Gaussian distribution on Dx was propagated to the reconstructed co-ordinates of node C. Node C was constrained strongly in X by the measurement CA and in Y by the measurement BC. The large but erroneous increase of Dx with temperature, therefore forced node C upwards in Z during the reconstruction to create the observed erroneously large correlation of the Cz with temperature. Node C also moved towards slightly towards the centre of the grid, reducing the X and Y co-ordinates, due to the slight tilt of the error ellipsoid of node C, previously discussed in the Monte Carlo test of Section 4.4.3.

The observations are therefore explained by the non-Gaussian measurement of GLI AB. The slight S-shape in the distribution of GLI AB with temperature is also faintly apparent in the reconstructed Z co-ordinate of node C. A correlation was found between these parameters as in Figure 4.30.

The stability test demonstrates that although the reconstructed co-ordinate did not exhibit the expected correlations with temperature, the reconstruction did proceed as expected from the available GLI data. Importantly, the recon-

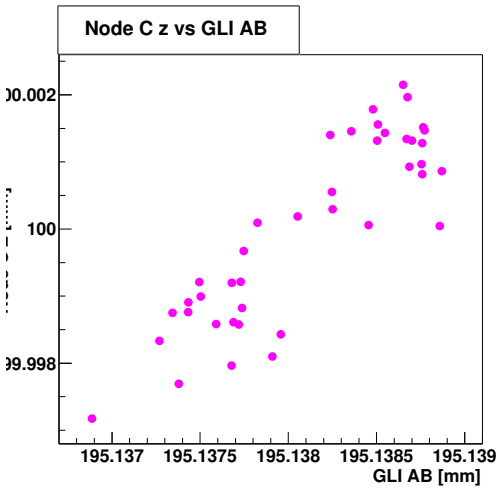

Figure 4.30: *Correlation of the reconstructed Z co-ordinate of node C with the measured length of GLI AB, for the tetrahedral grid.*

struction highlighted the error on the measurement of GLI AB, because the scatter on the reconstructed Z co-ordinate of node C was much larger than could be explained by thermal effects. The source of the increased measurement error on GLI AB was not obvious in further analysis of the interferometer data. Even when including this measurement error and the effects of thermal expansion, the precision on the three dimensional FSI grid reconstruction was found to be well within the ATLAS requirements.

4.4.6 Grid Self-consistency

The stability test data were reconstructed with a modified SIMULGEO model, in which node C was fixed rigidly in Z, so that only 5 DoF were solved for with the 6 GLIs. This over constrained model allowed a self-consistency check of the grid reconstruction, by systematically omitting GLIs from the analysis. The reconstructed co-ordinates for nodes B, C and D are plotted in Figures 4.31, 4.32 and 4.33 respectively for all analyses.

As found for the self-consistency check of the square grid, the mean values of the reconstructed node co-ordinates were independent of the combination of GLIs in the model, due to the appropriate calibration. The scatters on all reconstructed nodes were well within the precisions predicted when the required ATLAS FSI measurement precision was assumed, as in the second error propagation of Section 4.4.3.

The poor measurement of GLI AB for this test is apparent in the reconstructed X co-ordinate of node B. A non-Gaussian distribution was recon-

structed only for models in which AB was directly relied upon to determine the X co-ordinate of node B: when AB was omitted, the distribution had a more Gaussian shape. Notably, for the redundant model including all GLIs, the Bx distribution is not significantly affected by the poor measurement. The redundancy allowed the ratio of GLI measurement precisions to be accounted for in the adjustment procedure of SIMULGEO by assigning weights in accordance with the precisions. This reduced the adverse effects on the reconstruction of the non-Gaussian distribution of GLI AB. This emphasizes the importance of a redundant grid and the need for an appropriate calibration of the measurement precisions, for the best estimates on the node co-ordinates.

The precisions on each set of reconstructed node co-ordinates were calculated with Equation 4.2, for $N = 41$ FSI scans. The variation in the scatter was checked by an error propagation for each combination of 5 GLIs, assuming the length measurement precisions in Table 4.5. The predicted and reconstructed precisions on the node co-ordinates are compared for each model in Figure 4.34.

In general a close agreement was found between the predicted and the reconstructed precisions. The agreement to the nanometre for some node co-ordinates is explained as in Section 4.3.8.

The precisions include any effects of thermal expansion, particularly that induced by movement of node C in Z, which was assumed fixed in Z relative to node A. All precisions were well within the ATLAS requirements.

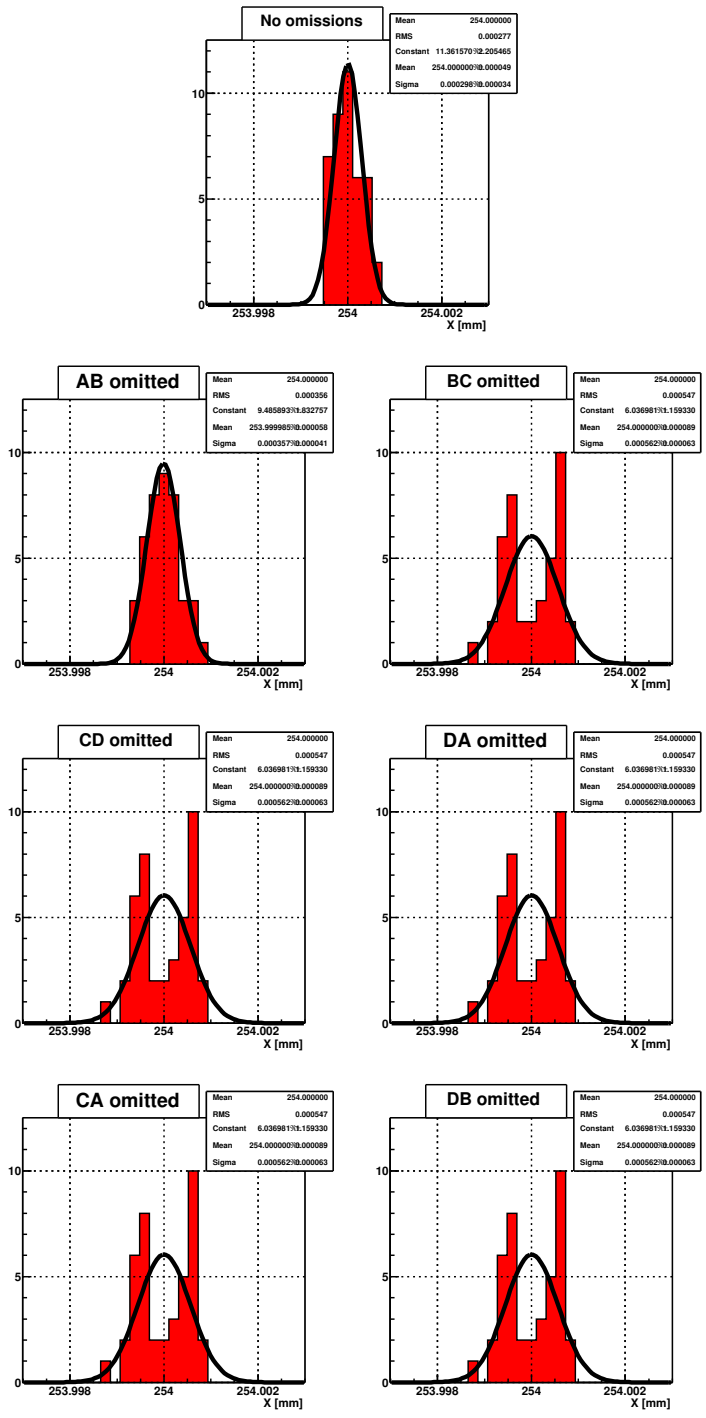

Figure 4.31: *The X co-ordinate of node B of the tetrahedral grid, assuming node C fixed in Z, reconstructed with models having the combination of GLIs that omitted the GLI stated in the plot title. The measured length of GLI AB had a non-Gaussian distribution.*

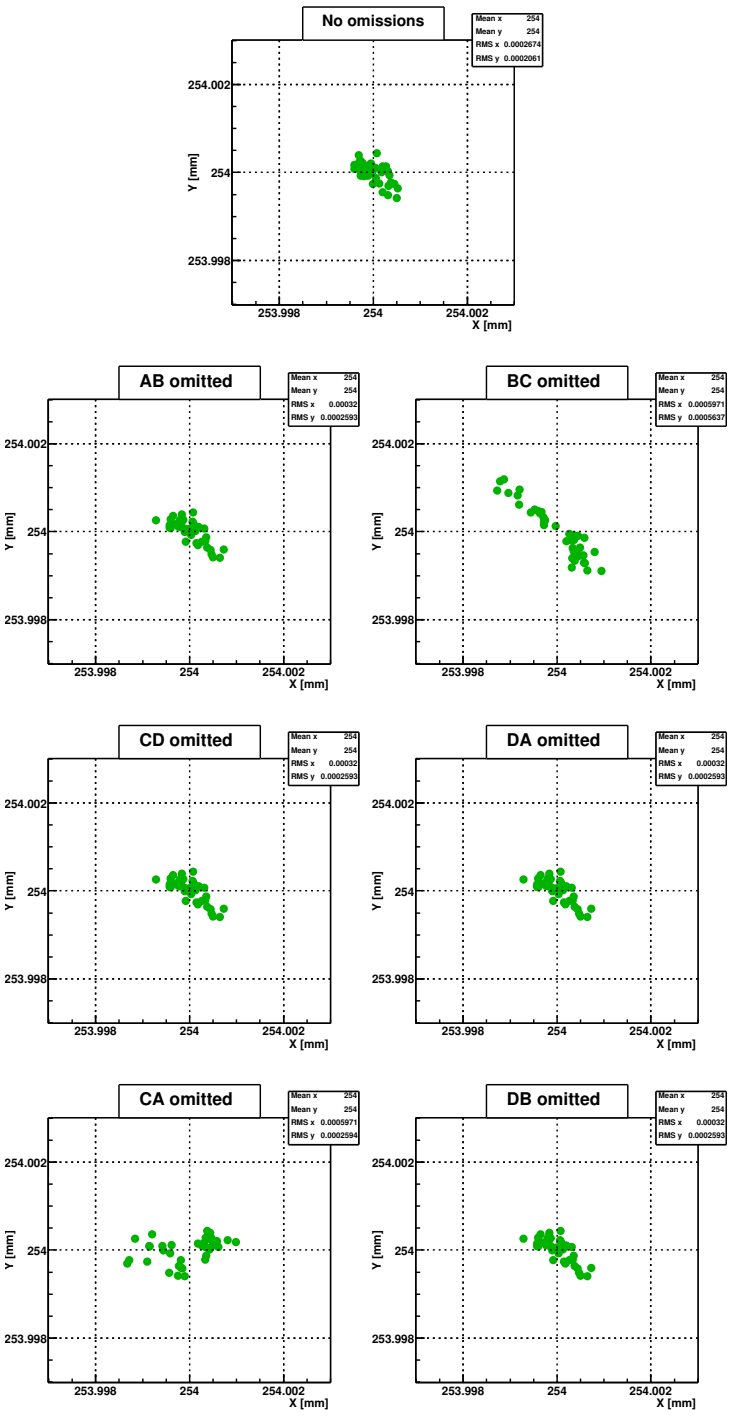

Figure 4.32: *The reconstructed X and Y co-ordinates of node C of the tetrahedral grid, assuming node C was fixed in Z. The GLI stated in the plot title was omitted in the analysis.*

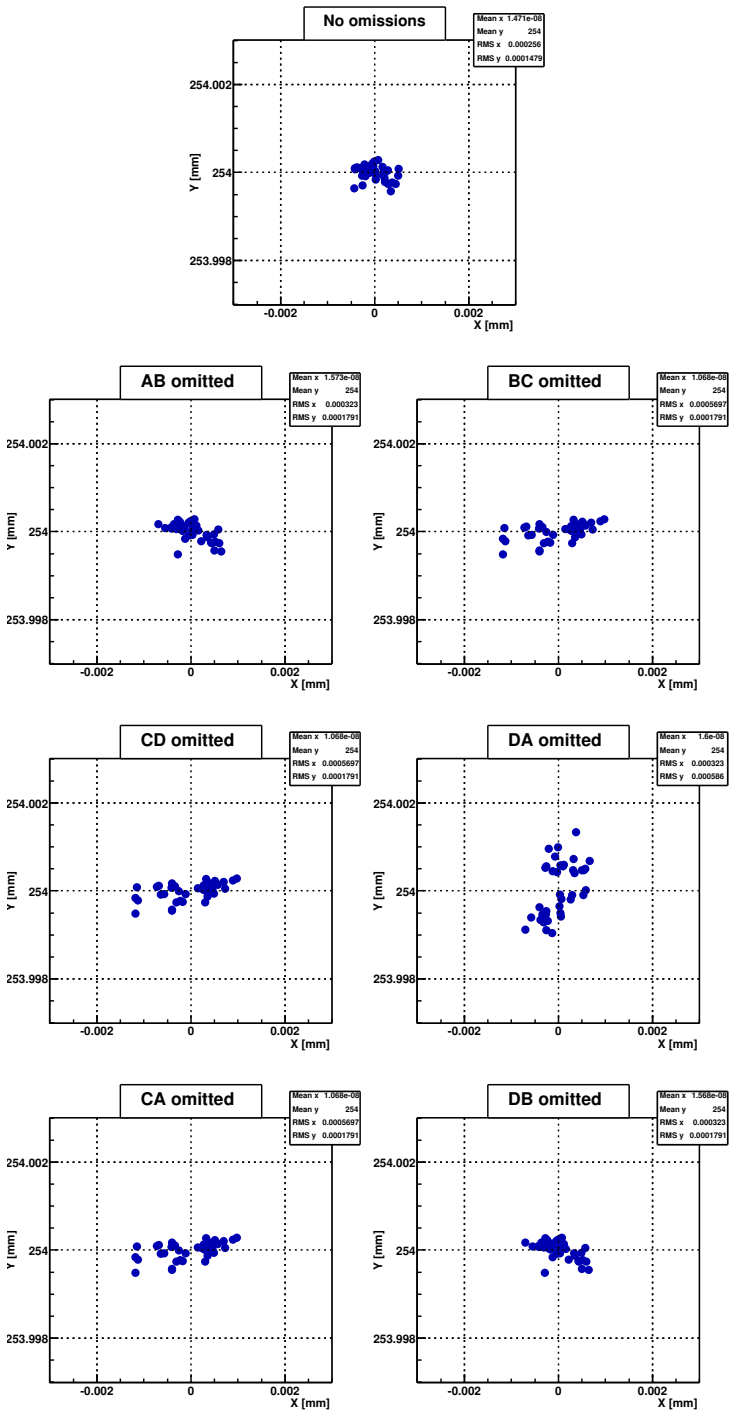

Figure 4.33: *The co-ordinates of node D of the tetrahedral grid, assuming node C fixed in Z, reconstructed with models having the combination of GLIs that omitted the GLI stated in the plot title.*

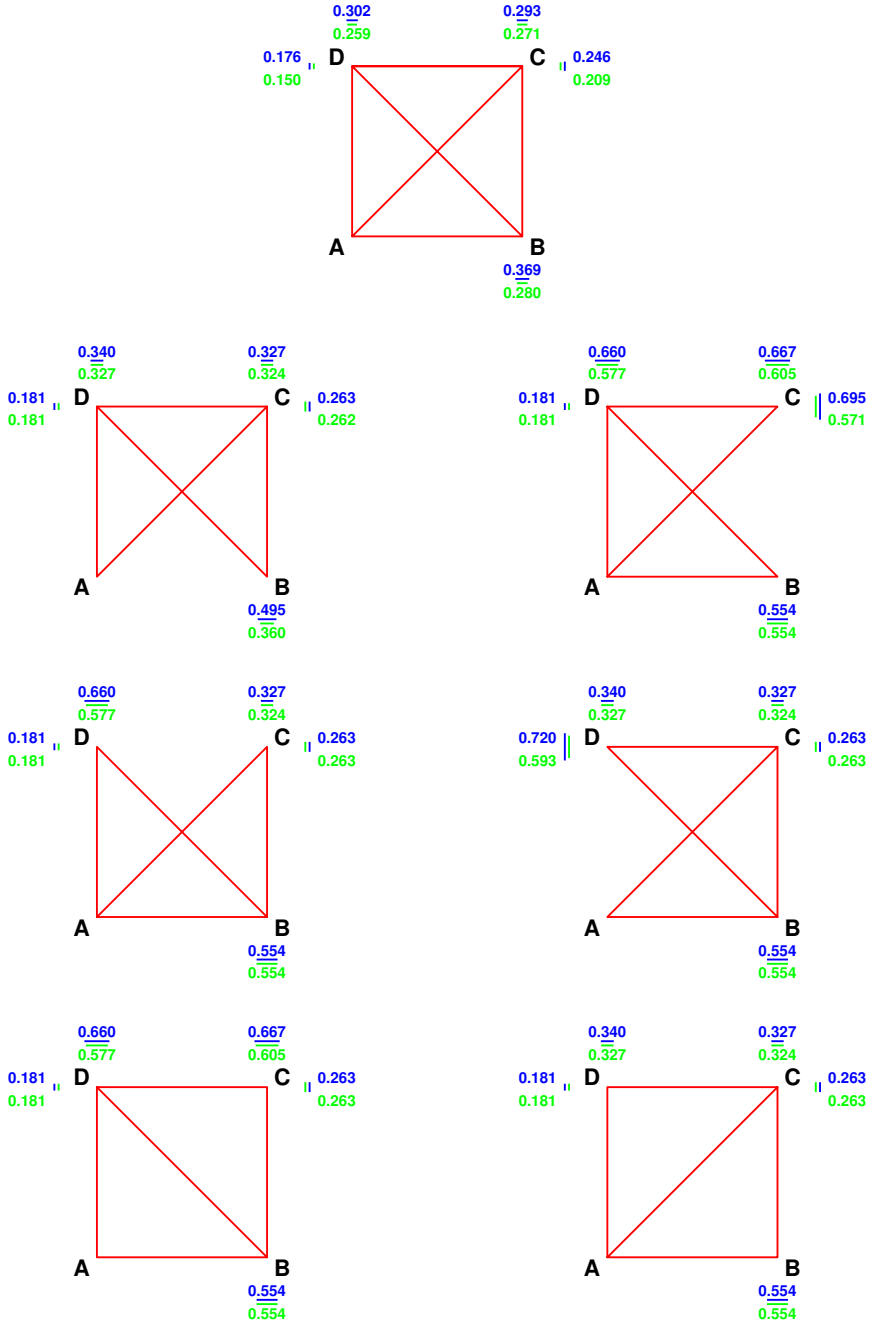

Figure 4.34: *Comparison of predicted and reconstructed precisions on the 5 free parameters of the tetrahedral grid, assuming node C fixed in Z, for each model that omitted the GLI as shown.*

4.4.7 Linearity Test

Measurement Procedure

The linearity on reconstructing the position of node C of the tetrahedral grid was evaluated as described in this section. 35 FSI scans were recorded as node C was displaced along the X-axis of the translation stage. 5 FSI scans were made at each of the 7 stage positions. The differential translation drive was turned once between each set of five scans, defining the step size of 47 μm ± 1 μm.

The measurements were recorded over 4.5 hours, during which time the temperature of the optical breadboard rose by 0.077 K.

Reconstruction

For each FSI scan the measured GLI lengths were extracted and the grid shape was reconstructed with the common calibrated SIMULGEO model, in which node C was free in Z. The reconstructed co-ordinates of node C are plotted in Figure 4.35. The 7 positions of the translation stage are apparent as the 7 groups of 5 points. The small reconstructed scatter of node C in the ZY plane indicates that the translation was approximately in the direction of the X-axis, defined as the direction between nodes A and B in model. The correlation of Z & Y shown in Figure 4.35 approximately matches the expectation from the Monte Carlo test shown in Figure 4.23.

The reconstructed X co-ordinate of node C was compared, as in Figure 4.36, with the stage displacement measured by interferometer X. The residuals from the best fit straight line were fitted by a Gaussian of width $(0.339 \pm 0.041)\mu$m, as plotted in Figure 4.37. The estimated reconstructed precision on node C in X was 0.300 μm from Figure 4.27. Therefore any non-linearity over the 0.28 mm translation range studied was unresolvable.

The gradient of the best fit straight line was smaller than unity by 0.0046, which is almost double the difference from unity observed in the square grid linearity test. This is most probably due to the further non-collinearity of the x-axis of the reconstruction model, the stage translation and the measurement direction of interferometer X, caused during the adaption to the tetrahedral geometry. Unfortunately, the design of the set-up did not allow the plate supporting the stage to be raised without dismantling the stage and interferometer X from the plate. Although every effort was made to ensure a consistent orientation of the stage and interferometer X direction, the difference observed would be caused by a further 0.5 mm relative shift in the ZY plane of nodes A or B.

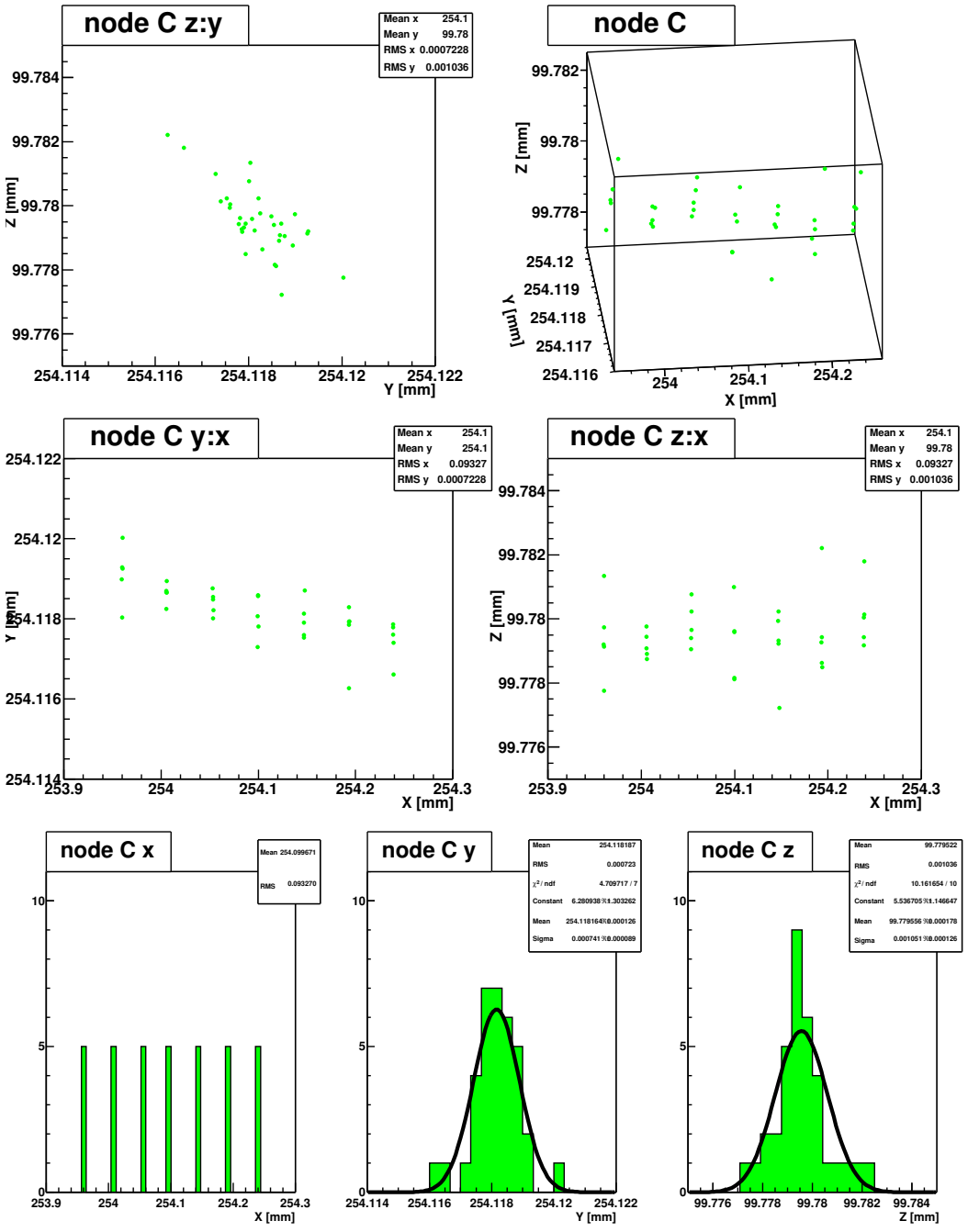

Figure 4.35: *The reconstructed co-ordinates of node C for the tetrahedral grid for the linearity test. The 7 positions of the stage can been seen.*

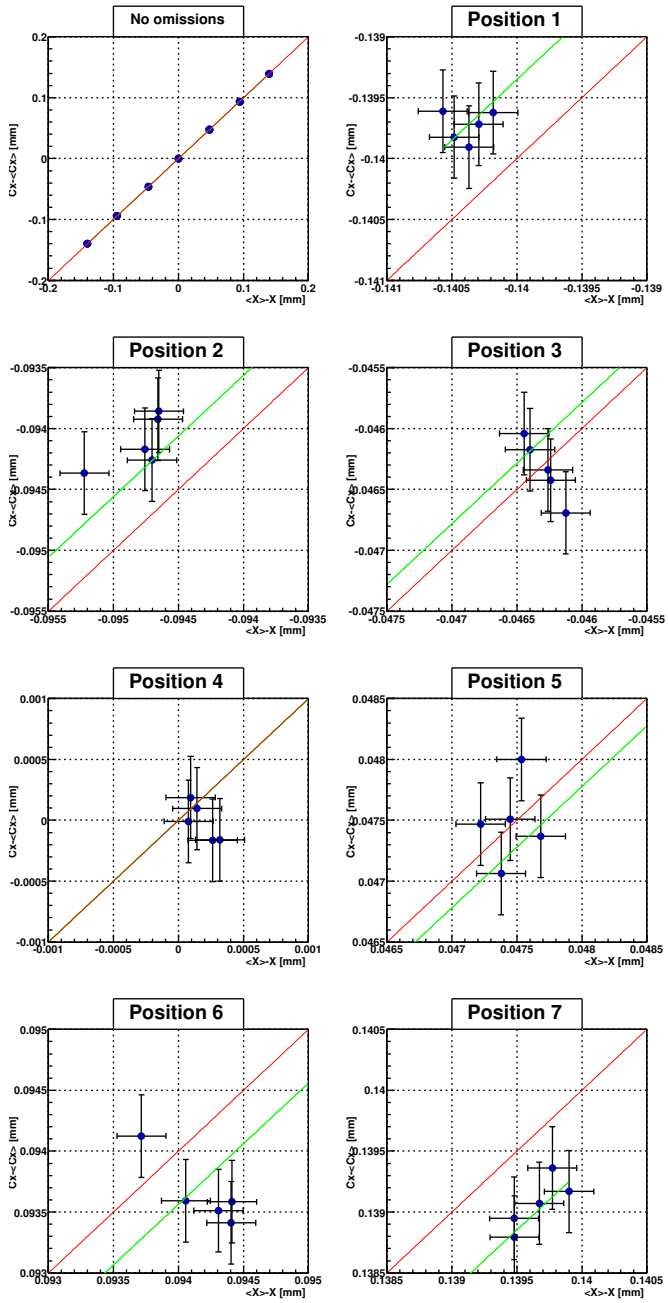

Figure 4.36: *The reconstructed X co-ordinate of node C for the tetrahedral grid versus the independently measured stage position, where all GLIs were used to reconstructed the grid. Magnified regions at the seven stage positions are plotted. The best fit straight line is shown in green (gradient 0.995360). A straight line with a gradient of 1 is shown in red for comparison.*

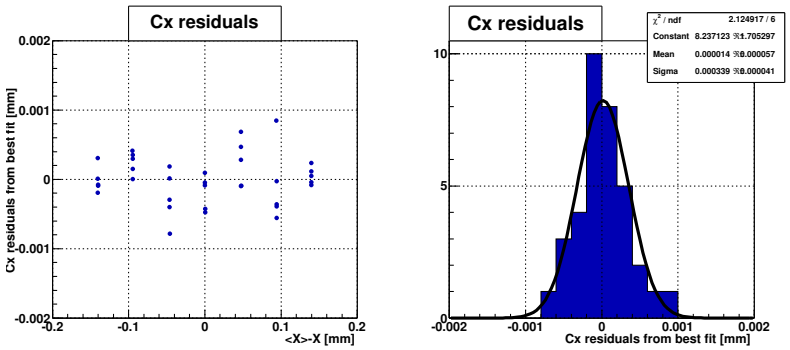

Figure 4.37: *The residuals of the reconstructed X co-ordinate of node C from the best fit straight line to the plot of Figure 4.36.*

4.4.8 Planarity Test

Measurement Procedure

The tetrahedral geometry allowed the reconstruction of displacements of node C in three dimensions to be investigated. The full extent of the X and Z differential axes were explored by recording ∼5 FSI scans at 9 different stage positions, which formed a 3×3 grid in the XZ plane. This allowed the planarity of the reconstruction to be evaluated. The step size of 141 μm ±2 μm between each stage position corresponded to 3 full turns of the X or Z differential translation drive.

The measurements were recorded over 6.3 hours, during which time the temperature of the optical breadboard rose by 0.259 K.

Reconstruction

The grid shape was reconstructed for each FSI scan with the common calibrated SIMULGEO model. The reconstructed co-ordinates of node C were found to closely follow the nine positions of the translation stage, as in Figure 4.38. The planarity of the reconstructed co-ordinates, is less than the fitted Gaussian width of 1 μm ±0.1 μm to the Y co-ordinate of node C. This is likely to be an overestimate due to the slight tilt in Y with increasing Z of the plane reconstructed co-ordinates that was observed, as shown in the ZY projection of Figure 4.38. The tilt angle of ∼0.4° is thought to be due to a misalignment in the set-up of the stage Z axis with the plane of grid nodes ABD, combined with the effects of thermal expansion.

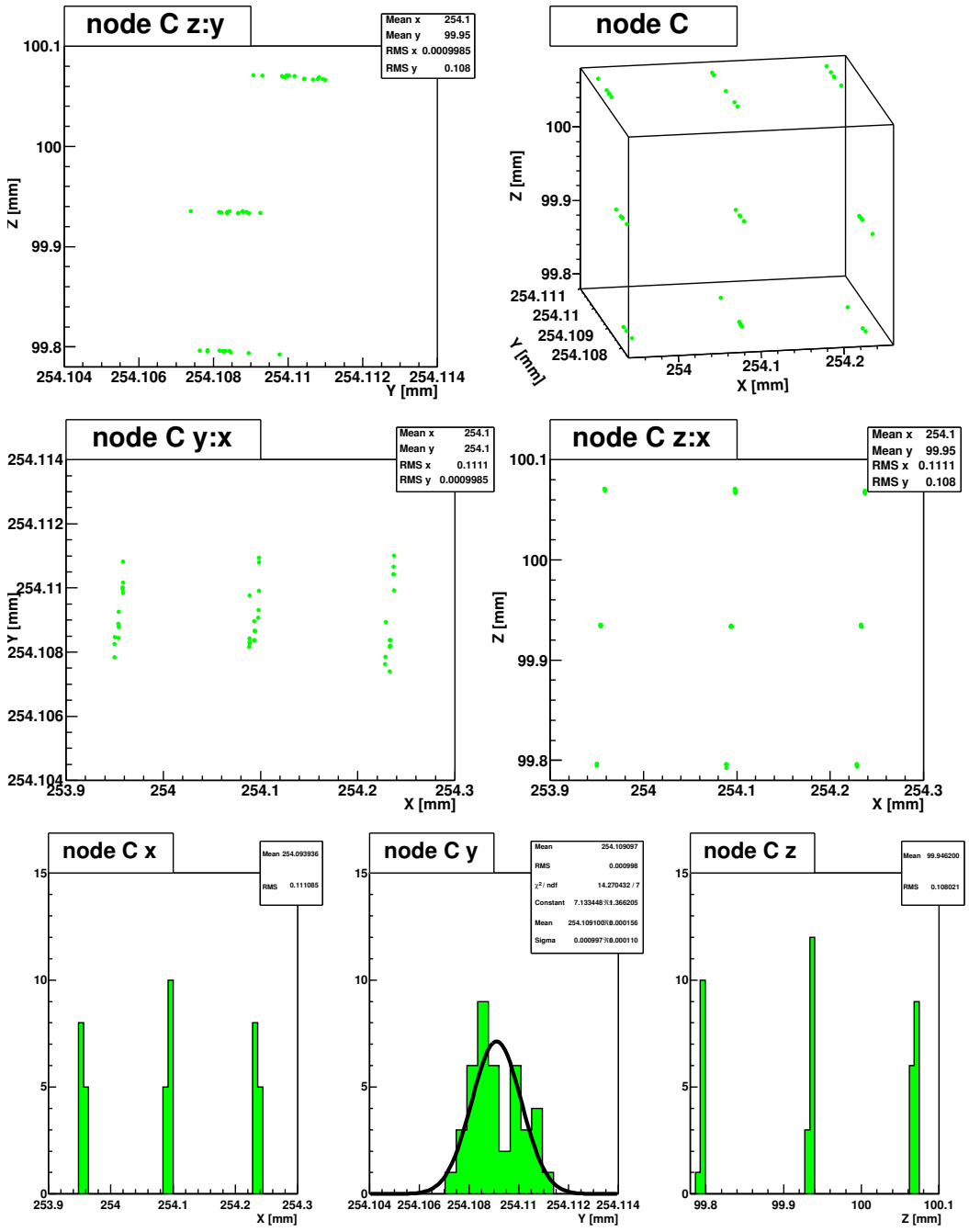

Figure 4.38: *The reconstructed co-ordinates of node C for the tetrahedral grid for the planarity test, showing the 9 stage positions that form a 3×3 array. A tilt of ∼0.4° in the ZY projection of the plane of reconstructed co-ordinates was observed.*

4.4.9 Predetermined Position Reconstruction

Measurement Procedure

In the final test with the tetrahedral grid, node C was sequentially moved in a complex pattern of predetermined positions and a single FSI scan was recorded at each stage position. The test aimed to demonstrate the typical role of FSI in ATLAS, in which a single FSI scan is required to precisely determine a particular SCT shape, in case the shape changes are too fast for the precision to be improved by multiple FSI scans.

The position of node C was manipulated with the X and Z differential translation drives of the stage, in quarter turn intervals. The step size between adjacent node positions was 12 μm $\pm$1 μm, which is equivalent to the RΦ precision required for the ATLAS FSI system.

The measurements were recorded over 6.7 hours, during which time the temperature of the optical breadboard rose by 0.412 K. This was the largest temperature change for any grid test, due to the repeated opening of the thermal insulation to adjust the stage, between each $\sim$8 minute FSI scan.

Reconstruction

The grid shape was reconstructed for each FSI scan with the common calibrated SIMULGEO model. The reconstructed co-ordinates of node C followed closely the 39 positions of the stage, as plotted in Figure 4.39.

The reconstructed position of node C was found to lie within the XZ plane, to within 0.80 μm $\pm$0.09 μm calculated from the width of a Gaussian fit to the reconstructed Y co-ordinate of node C. A tilt in Y with increasing Z that was found, that was consistent with the 0.4° observed in the planarity test. The spread in Y was less than for the planarity test, because of the tilt and reduced range in Z of the measurements.

Importantly, the FSI grid shape measurements precisely determined the complex pattern of node movements to well within the ATLAS requirements, as shown in Figure 4.40. The precision was demonstrated by taking the residual differences between the reconstructed co-ordinates and a set of expected co-ordinates based on a quantized constant separation of the node positions, as in Figure 4.41. The scatter in these residuals is dominated by the effective stage setting sensitivity, rather than the FSI measurement precision, which is shown for stability test data in Figure 4.42 for comparison.

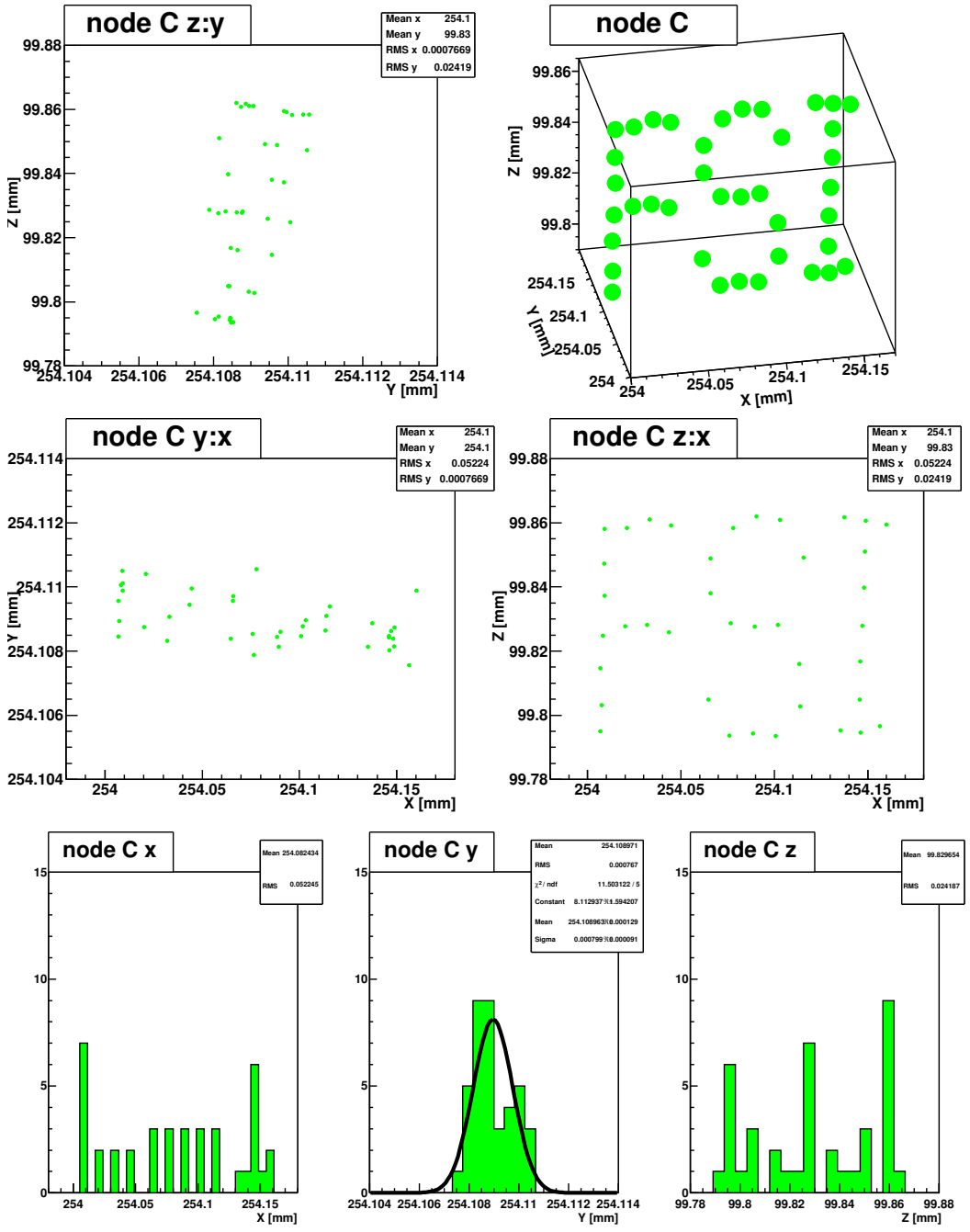

Figure 4.39: *The reconstructed co-ordinates of node C for the tetrahedral grid. The predetermined complex pattern of 39 stage positions was precisely reconstructed by FSI grid measurements. A tilt of ∼0.4° in the ZY projection of the plane of reconstructed co-ordinates was observed.*

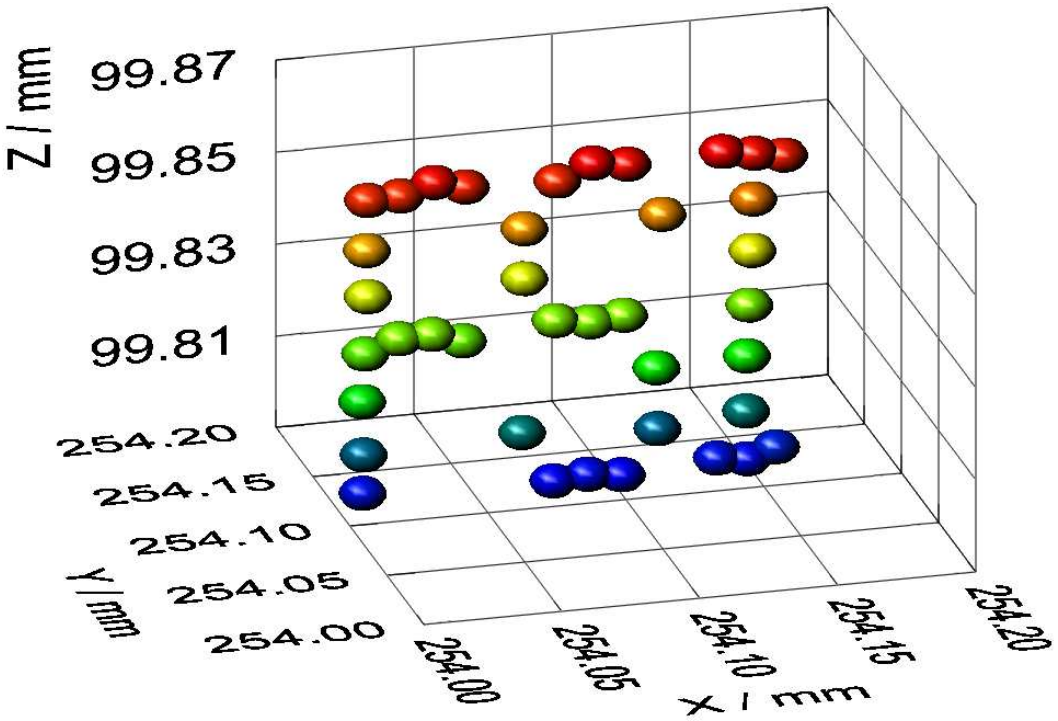

Figure 4.40: *The reconstructed co-ordinates of node C of the tetrahedral grid precisely followed the complex pattern of 39 node positions to which node C was sequentially set. The positions are clearly distinguishable, despite the small separation between adjacent points of 12 µm, corresponding to the required RΦ precision for the ATLAS FSI system.*

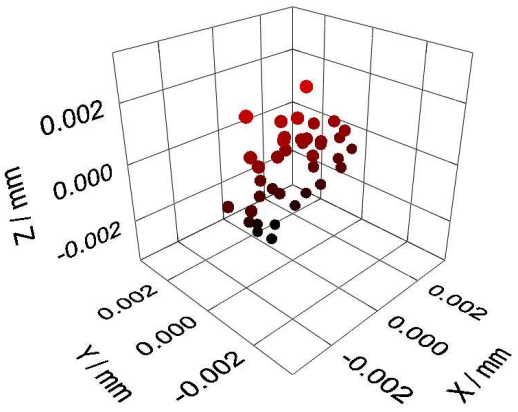

Figure 4.41: *The residuals of the positions in Figure 4.40, and the expectations from perfect stage movements. The scatter was dominated by the stage setting sensitivity, rather than the FSI precision, in Fig. 4.42.*

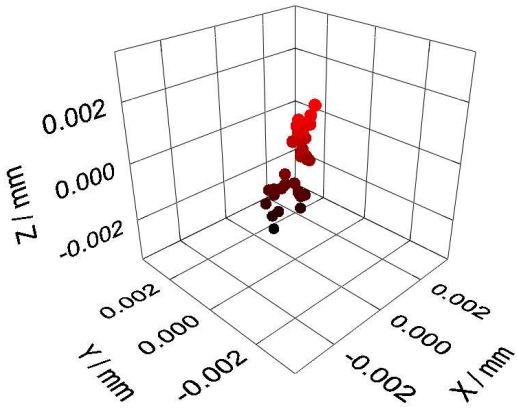

Figure 4.42: *The scatter in reconstructed positions of node C for the stability test, in which the stage supporting node C was not adjusted. The scatter is due to the FSI precision (and the thermal expansion of the set-up).*

4.4.10 Discussion and Conclusions

Three dimensional FSI geodetic grid shape measurement has been demonstrated for the first time with the tetrahedral grid. The precisions achieved on the reconstructed tetrahedral grid node co-ordinates were on average a factor of 4.3 better that the precision required for ATLAS.

A method for calibrating the software model of a three-dimensional FSI grid has been developed, that could be extended to aid calibration of the ATLAS FSI system. The grid shape was successfully reconstructed with the calibrated model, despite the removal of any one GLI from the analysis, when node C was assumed fixed in Z.

The correlations between the node co-ordinates reconstructed from FSI grid measurements were consistent with the expectations from simulated data. An analysis of correlations of the reconstructed node positions and GLI lengths with temperature highlighted the poor measurement on GLI AB.

The deliberate displacements of one grid node were precisely followed in the reconstruction of FSI grid measurements. Any non-linearity or non-planarity of the reconstructed co-ordinates were less than the FSI measurement precision.

Complex patterns of node movements were precisely reconstructed from a single FSI grid measurement at each node position, as required for the ATLAS FSI system.

4.5 Conclusion

Small shape deformations of the first two and three dimensional FSI geodetic grids have been precisely reconstructed from simultaneous FSI length measurements between grid nodes. The precisions on the node co-ordinates were found to be on average a factor of $\gtrsim$4 better than the precision required for the ATLAS FSI system.

Reconstruction models were developed to allow the shape of the prototype grids to be determined from FSI measurements. An important calibration technique was developed, that could be extended to the ATLAS FSI system. An appropriate calibration was crucial for the grid shape to be reconstructed consistently when redundant grid lines were removed.

The precisions on the reconstructed node co-ordinates improved for a redundant grid. The overconstrained grid allowed the poorer measurements to be assigned less weight in the adjustment procedure, resulting in a better estimate of the node positions.

Correlations between the reconstructed node co-ordinates were as expected from simulated data. Such correlations could be important in defining an ap-

propriate set of degrees of freedom for the ATLAS FSI grids.

Each of the simultaneous FSI measurements were in general found to be very precise ($\sim$ 200 nm), despite any remaining optical cross-talk between GLIs in the set-up. For one GLI the measurement precision was found to be particularly poor ($\sim \pm 554$ nm), although the reconstruction was tolerant to this in the case of a redundant grid. This emphasizes the need for estimated GLI measurement precisions to be included in the reconstruction models for the ATLAS grids.

The simulated predictions of rigidity of the prototype grids closely matched those calculated from the reconstructed grid shapes for repeated grid measurements, inspiring confidence in the simulated predictions on the rigidity of the ATLAS FSI grids described in Chapter 5.

Chapter 5

ATLAS FSI Grid Simulations

5.1 Introduction

The prototype grids of Chapters 3 and 4 are very simple in comparison to the ATLAS SCT alignment grid, which consists of 842 measurement lines arranged in a complex geometry. The ATLAS grid is divided into three sections that simultaneously monitor the SCT barrel (512 lines) and the two SCT end-caps (165 lines each).[1] Such a vast grid could not be fully prototyped, due to limited resources and available power from the prototype FSI system. The expected precision in reconstructing the ATLAS grid has been calculated with detailed computer models, written in the same simulation software as used to successfully reconstruct the prototype grids (see Chapter 4).

The presented studies are the first simulations of the final design of the ATLAS FSI grid. This chapter summarizes detailed investigations of the propagation of errors through the various sections of the grid. The studies aimed at understanding the performance of the grid and determining if and how, various regions of the grid relate to each other, during the reconstruction process. The knowledge gained should enable the node reconstruction method to be optimized for ATLAS alignment.

The studies address how each grid section has been configured to fulfill the ATLAS alignment requirements, which are reviewed in Section 5.2. The SCT Barrel and SCT End-cap grid sections are evaluated in Sections 5.3 and 5.4 respectively. The implementation of the FSI system and how it could be integrated with track alignment is outlined in Section 5.5.

[1] Relative alignment of the SCT barrel and end-caps will be performed with tracks.

5.2 ATLAS and FSI Grid Design

5.2.1 ATLAS Alignment Grid Requirements

To ensure that misalignments do not degrade any track parameter resolution by more than 20%, the SCT detector modules should be aligned with a precision in RΦ of 12 μm [ATL97, Wei97].[2] The less stringent corresponding Z and R alignment requirements, of 50 μm and 100 μm respectively for the SCT barrel, should be easily met by an FSI grid that fulfills the RΦ requirement, as shown in previous studies [Fox96]. The requirement has been re-expressed as a precision of around 10 μm on the three dimensional reconstruction of the FSI grid nodes [Nic96]. However, the SCT module co-ordinates will be interpolated from the FSI grid node positions and refined with track data, as discussed in Section 5.5. Interpolation errors are not yet known, but it has been suggested that to achieve the required precision on the module co-ordinates, the RΦ precision of the alignment grid should be around 5 μm [Fox96].

Expected SCT deformations

The purpose of the alignment grid is to determine deformations of the SCT. Several potential deformations have been investigated with FEA models [Cug01], including those induced by movements of structures attached to the SCT, namely the SCT support rails and the pixel detector support tube. Importantly, certain deformations are invisible to track based alignment methods, which rely on the deviation of the track residuals from a helical path. These deformations cause tracks to remain helical but change one or more track parameters (e.g. curvature) and are impossible to correct in some cases, with track data alone [Hin04]. Certain classes of deformation may be constrained by comparing the track momentum with the energy measured in the calorimeters, for oppositely charged tracks. However this method is subject to complicated systematic effects because the numerous track statistics required can only be accumulated over time-scales that may exceed those of the deformations, and the method cannot constrain all deformation classes. Therefore, an important requirement is that the alignment grid is sensitive to all deformations that are not well constrained by tracks (such as *sagitta deformations*). Considering the deformation of the SCT as being decomposed into a set of eigenmodes, then the FSI system should be sensitive to the low spatial frequency eigenmodes, which are the least accessible with track data.

[2] An RΦ alignment of 1 μm is estimated for the M_W fitting requirement [Hay00]. This aim may be achievable with combined track and FSI alignment methods.

5.2.2 Geodetic Grid Design

For a geodetic grid to be over-constrained, the sum of measurements and imposed physical constraints must be greater than the number of degrees of freedom (DoF) of the grid nodes. A well designed grid can then be solved to allow reconstruction of the grid node positions and the orientations of extended objects.

As demonstrated in Chapter 4, the precision in reconstructing the grid node co-ordinates is strongly dependent on the number and arrangement of the grid line interferometers (GLIs). Without spatial constraints, a rigid grid may be configured with relatively few GLIs, by arranging them into rigid shapes such as equilateral triangles and braced-quadrilaterals. In the SCT however, mechanical boundaries confine the GLIs to compact regions between module layers, services and support structures. The number of GLIs required to achieve adequate grid rigidity increases in regions with awkward spatial constraints.

Redundancy in the grid improves the grid rigidity, as shown with the prototype grids in Chapter 4 and by definition ensures that the grid can still be reconstructed in the case of redundant GLI measurement failure. However, the number of GLIs is limited principally by the power of the laser system and cost.

The ATLAS alignment grid has been configured for rigidity and redundancy, within the restrictions outlined above. Further information on geodetic grid design can be found elsewhere [Fox96, Kua96].

5.3 SCT Barrel Grid Simulations

This section presents detailed studies of error propagation through the SCT Barrel grid. Previous investigations of the SCT barrel grid indicated that the ATLAS requirements could be met with a geodetic grid similar to those studied earlier [Fox96]. The preliminary grid designs were developed into the final design [How01], which has subsequently been evaluated in more detail, as described here.

5.3.1 SCT Barrel Grid Design

The ATLAS SCT barrel consists of four concentric carbon-fibre cylinders supporting the silicon strip detector modules. Carbon-fibre flanges are attached at the ends of each cylinder to increase rigidity. The cylinders are joined together by eight radial interlink beams at each flange. The cylinders are labelled as barrel 3 to barrel 6, with increasing radius.

The SCT barrel alignment system is separated by mechanical boundaries into four inner barrel grids (one per barrel) and two barrel-flange grids. The grids are formed between FSI jewels that are rigidly attached to the carbon-fibre support structures.

Each inner barrel grid monitors the shape of the barrel to which it is attached. The inner barrel grid is constricted to a small annular region in the RΦ plane, having the radial width of the flange from the inner surface of the barrel. The grid comprises 32 grid nodes connected by 80 GLIs, configured as 8 unit cells that each survey an octant of the barrel. The GLIs of each octant run approximately parallel to the barrel axis. Two additional GLIs from each octant measure radially inwards to a retro-reflector located on a module of the adjacent smaller barrel[3], to help constrain the radial deformations of the barrels.

The two barrel-flange grids monitor the relative positions and orientations of the four barrels. The grid measurements link the barrels in the transverse RΦ plane at either SCT barrel flange. Each barrel-flange grid comprises 72 GLIs divided into eight sectors, situated between the radial interlink beams. The measurements also help to constrain deformations of the end-flanges of each barrel.

5.3.2 Overview of Barrel Grid Simulations

The SCT barrel alignment system must combine information from the 512 GLIs forming the grid sections described above, to determine the shapes and relative positions and orientations of the barrels. This task was broken down and studied in the following phases:

- inner barrel grid
  - (a) single barrel octant; the unit cell (11 GLIs, including 1 GLI from adjacent octant)
  - (b) single barrel quadrant (21 GLIs, including 1 GLI from adjacent quadrant)
  - (c) complete inner grid of a single barrel (96 GLIs; 80 GLIs, excluding radial lines)
- barrel-flange grid linking one end of all barrels (72 GLIs)
- complete SCT barrel grid (512 GLIs; 464 GLIs, excluding radial lines)

[3]The inner grid of the smallest SCT barrel does not have these radial GLIs.

The aim was to elucidate the rigidity and redundancy of each grid section and aid understanding of the complete barrel grid. Several progressively complex simulations were performed in each phase, with models written in the SIMULGEO [Bru98] software described in Section 4.2.4. The precision of the grid was evaluated with the error propagation mode of the software. The simulation method differed from Chapter 4 in that some grid nodes were locally constrained with particular weights, as detailed in each simulation section, rather than fixed absolutely. The weights were accounted for in the adjustment procedure, giving a more realistic estimate of the grid precision.

The SCT barrel grid is not over-constrained solely by length measurements, but also relies on physical constraints imposed by the allowed deformations of the carbon-fibre support structure. Nothing was assumed regarding correlated movements of FSI nodes induced by certain modes of barrel deformation. Rather, where necessary, each jewel was constrained simply by tolerances estimated with FEA models [Cug01].

5.3.3 Inner Barrel Grid

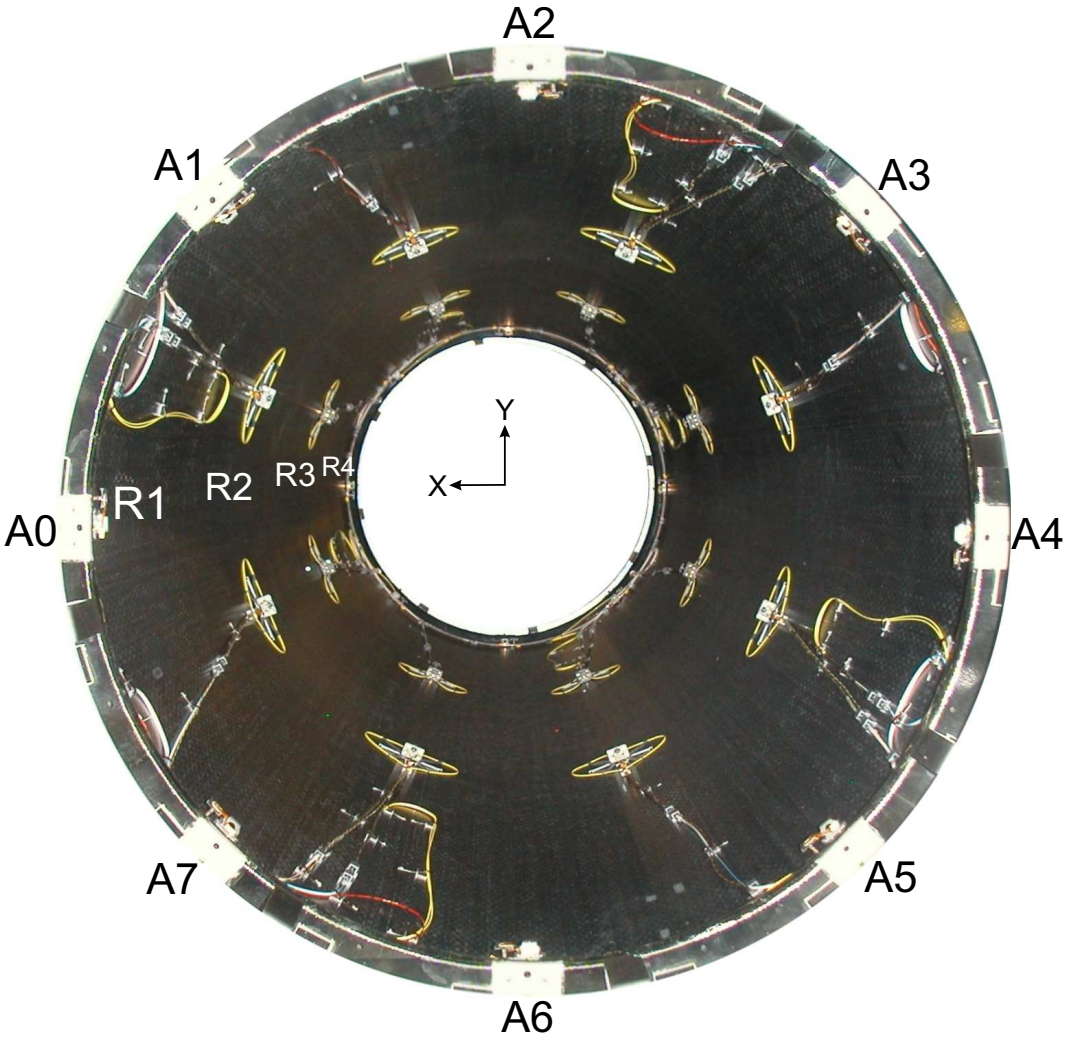

Figure 5.1: *The ATLAS SCT barrel 3 alignment grid. The interior surface of the Ø 560 mm carbon-fibre cylinder is shown as photographed from the beam axis. 80 Grid Line Interferometers are formed between 32 FSI jewels. The jewels are rigidly attached to the carbon-fibre, forming 4 rings, R1-R4, that are approximately equidistant in Z along the 1.5 m barrel length. Rings R2 and R3 consist entirely of jewels called scorpions that hold the majority of quills on the barrel. Each scorpion can be seen between a curving pair of yellow furcation tubes, which protect the optical fibre powering the quills. Rings R1 and R4 consist entirely of endjewels that hold mainly retro-reflectors and are attached close to the rigid barrel-flanges. The barrel jewels are the grid nodes, labelled by barrel B3-B6, ring R1-R4 and azimuth A0-A7.*

The inner alignment grid for barrel 3 of the ATLAS SCT is shown in Figure 5.1.

(a) Single Barrel Octant Grid

An isolated octant of the barrel grid was simulated to evaluate the precision in reconstructing the octant grid, without the additional redundancy of adjacent octants. The intent was to aid understanding of the complete grid and verify that one octant grid meets the ATLAS requirements. This verification would ensure that the alignment system locally meets the specification even if measurements connecting grid octants fail.

One octant of a single barrel grid consists of 6 FSI jewels connected by 11 GLIs, as shown in Figure 5.2. The octant is configured as two adjacent braced-quadrilaterals that each resemble the prototype square grid of Chapter 4, distorted into a trapezium. The FSI jewels represented by circles in Figure 5.2 are attached close to the two rigid carbon-fibre flanges.

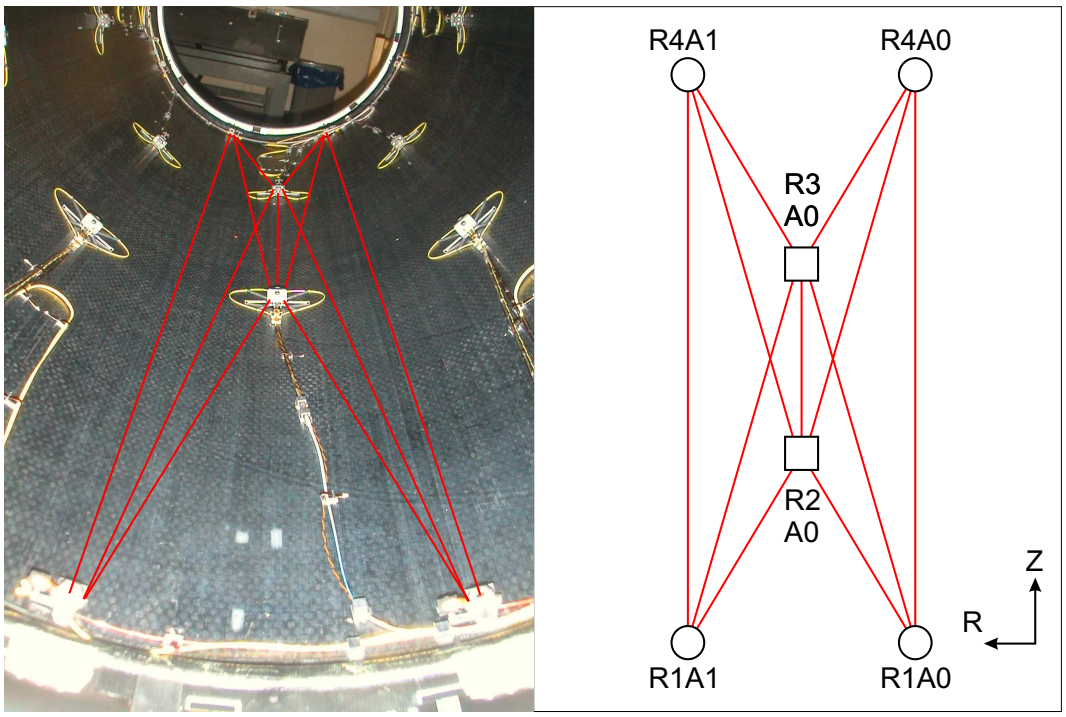

Figure 5.2: *One octant of the inner alignment grid for a single SCT barrel. The alignment system of SCT barrel 3 is shown on the left with the GLIs of one octant overlaid with red lines. The octant grid is shown schematically with the node labels on the right. The grid spans the 1.5 m barrel length.*

A model of the octant grid was written in SIMULGEO. The grid dimensions and the relative displacements of the quill and retro-reflector nodes within the jewels were taken directly from engineering drawings [Bar01, Bro02]. No relative movement was allowed between the nodes within a jewel.

The model required a reference co-ordinate frame to be defined. Three non-collinear points are sufficient to define a co-ordinate frame that has 6 external DoF and 3 internal DoF. For the octant grid simulation, the three grid nodes R1A0, R1A1 and R4A0 were chosen to define the co-ordinate frame, due to their large separation. Node R1A0 defined the origin and the 3 internal DoF were defined as the RΦ co-ordinate of node R1A1 and the Z and RΦ co-ordinates of node R4A0. The precision on reconstructing all other grid nodes was determined with respect to this reference frame.

A completely free simulation of the octant grid does not solve independently. Even when neglecting jewel rotations, the 6 FSI jewels have 18 translational DoF, 6 of which are taken up in the definition of the reference frame. The 12 remaining DoF cannot be solved with just 11 GLI measurements.[4]

Substantially more GLIs would be needed to solve for all translational and rotational DoF of the jewels. Instead, the grid relies on additional constraints imposed by the rigidity of the carbon-fibre support structure. In particular, the rotations of each jewel about the local axes should be tightly constrained. These local rotations were fixed almost completely (to 0.1 μrad) in the simulations. This simple simulation omitted the radial measurements from the adjacent barrel. The radial co-ordinates of the jewels were instead fixed to 5 μm. Initially rigid barrel-flanges were not assumed. All jewels not defining the reference frame, were free to move in RΦ and Z.

The grid was solved in three dimensions, assuming a 1 μm measurement precision for all GLIs. The simulated precisions in determining the RΦ and Z co-ordinates of each grid node are plotted against the Z location of the grid node in Figure 5.3.

The pattern of simulated precisions is peculiar to the DoF definitions in the model. In general, the RΦ precision worsens with the Z location of the grid node from Ring 1. This is because Ring 1 contained the reference node, R1A0, that defined the RΦ reference and also because the nodes R1A0 and R1A1 defined the baseline for the RΦ plane. The worsening of the RΦ precision with distance from this baseline is comparable to the increase in lateral sway of a tall structure, such as a tower block, with height from its base. This analogy also explains the stark contrast in magnitude of the RΦ and Z precisions in Figure 5.3: the lateral sway is much larger than the extension or compression

[4]The remaining DoF can be imagined as the grid folding about the line through nodes R2A0 and R3A0.

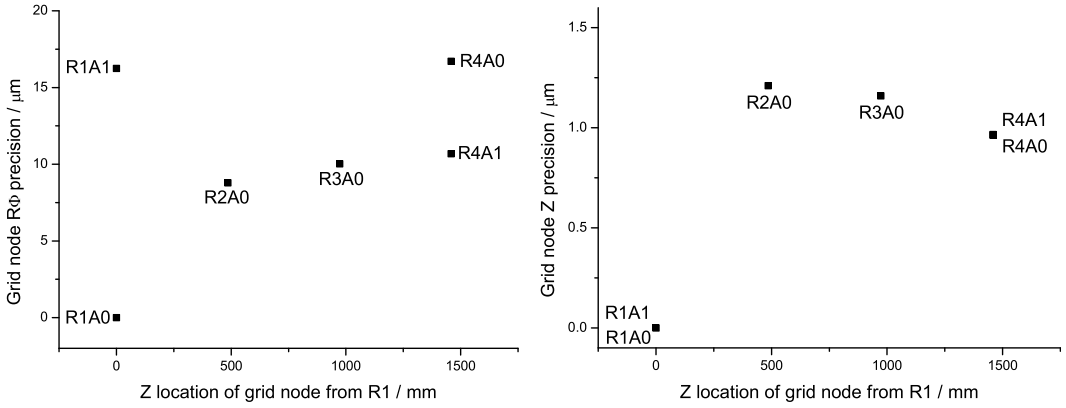

Figure 5.3: *The RΦ and Z precisions on the grid nodes of an octant grid of Barrel 3, simulated without assuming rigid barrel-flanges. The precision is plotted versus the grid node Z location from Ring 1.*

of the tower block height.

The Ring 4 node Z precisions are better than the measurement precision because of the over-constraints created by direct in-Z measurements and indirect measurements from the baseline. The GLIs connecting the two central nodes to the baseline are angled away from the Z direction, reducing the Z component of the measurements and worsening the Z precision.

The above simulation demonstrates the precision achievable with a single octant grid, when nothing is assumed about the relative separation of the endjewels within a ring. In this case the RΦ precision clearly does not meet the ATLAS specifications. However in the SCT, the endjewels are attached deliberately close to the rigid carbon-fibre flanges. Also, FSI nodes situated near the endjewels, on the opposite side of the flange, are monitored by the barrel-flange grid. The relative RΦ position of adjacent endjewels within a ring should therefore be known to better than 2 μm [Lau01].

The simulation was modified to include the 2 μm RΦ constraint. The endjewels of Ring 4 were defined in a distinct co-ordinate frame that was free to translate in Z and RΦ of R4A0 and rotate about the local R-axis of R4A0. The precisions were evaluated repeatedly for each barrel radius[5] and are compared in Figure 5.4.

A clear improvement in both the RΦ and Z precisions is observed when the rigidity of the barrel-flange is included in the model, as shown by the Barrel 3 results in Figure 5.4 compared with Figure 5.3. The two extra constraints

[5]For barrels 4 and 6 a small additional azimuthal offset of the central jewels exists in the grid design, but was excluded from the simulations to simplify the comparison.

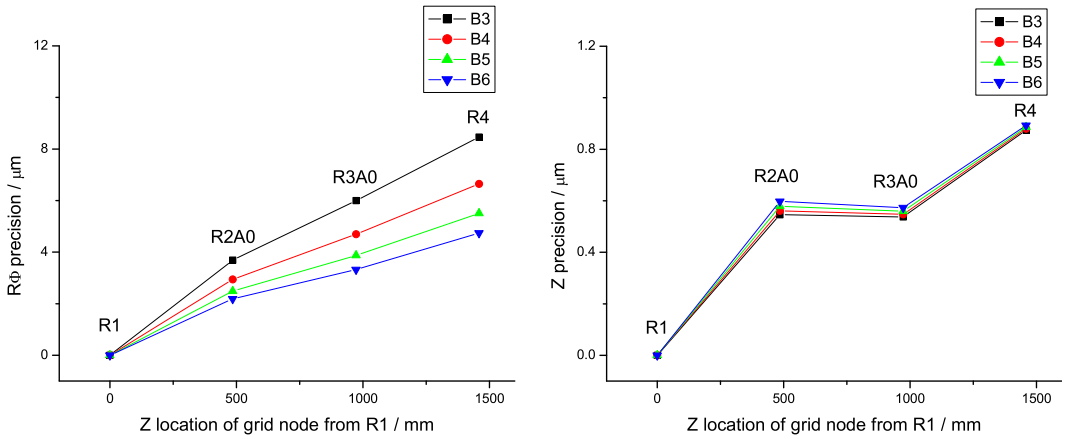

Figure 5.4: *The RΦ and Z precisions on the grid nodes of an octant grid for each Barrel 3-6. The rigidity of the barrel-flanges was modelled. The precision is plotted as a function of the grid node Z location from Ring 1.*

reduce the number of DoF to be solved, increasing the grid redundancy and hence rigidity.

On comparing the grid node precisions for different barrels, the RΦ precision was found to improve and the Z precision to worsen, as the barrel radius increased. The barrels have a common length and grid arrangement. Therefore the length:width aspect ratio of the octant grid reduces with radius; from 7:1 for Barrel 3 to 4:1 for Barrel 6. As the barrel radius increases, an RΦ component increase and Z component decrease occurs for GLIs connecting the endjewels with the two central jewels, producing the observed consequences on the grid rigidity. The RΦ component changes faster than the Z component due to the geometry, so the effect is larger for the RΦ precision than for Z.

The octant grid simulation has shown that valuable alignment information may be gained, without relying on assumptions regarding the barrel-flange. When the rigidity of the barrel-flange is modelled, an isolated octant grid meets the ATLAS requirements, with the jewel rotations constrained.

(b) Single Barrel Quadrant Grid

The barrel octant grid simulation was extended to a quadrant of the barrel, to examine how the additional redundancy improved the precision in reconstructing certain nodes. The quadrant grid contains four additional jewels (R1A2, R2A1, R3A1 and R2A2) making a total of 10 jewels linked by 21 GLIs.

The definition of the reference co-ordinate frame was identical to that in the octant grid model, to allow a direct comparison. The additional jewels were

free in Z and RΦ and were fixed to 5 μm in R.

Initially, nothing was assumed regarding the barrel-flange rigidity. The precision on determining the RΦ co-ordinate of a jewel in each ring is plotted against the Z co-ordinate of the jewel in Figure 5.5.

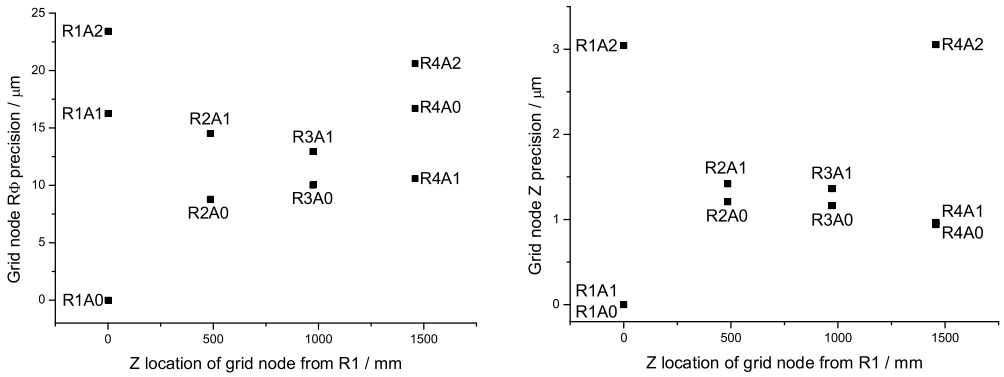

Figure 5.5: *The RΦ and Z precisions on the grid nodes of a quadrant grid of Barrel 3, simulated without assuming rigid barrel-flanges. The precision is plotted versus the grid node Z location from Ring 1.*

A comparison of Figure 5.5 and the corresponding Figure (5.3) for the octant, reveals no significant improvement in the RΦ or Z precision on any jewel of the original octant grid: the same pattern of precisions is observed. This important null result shows that without barrel-flange assumptions, the addition of an adjacent octant does not significantly improve the octant grid rigidity.

As with the octant, the simulation was modified to model the rigidity of the barrel-flanges. Ring 1 was the reference and Ring 4 was defined in a distinct co-ordinate frame that was free to translate in Z and rotate about the local R-axis of R4A0 and also about the Z-axis of the barrel, to simulate coupled RΦ movement of the jewels constituting Ring 4. Nodes R1A2 and R4A2 were fixed locally within the relevant rings to 5 μm in R, $2\sqrt{2}$ μm in RΦ and 2 μm in Z [Lau01].

The simulation was performed for each barrel radius and the precisions are compared in Figure 5.6.

As with the octant grid, a clear improvement is observed in both the RΦ and Z precisions for each node of the quadrant grid, when the rigidity of the barrel-flange is modelled, as shown by the Barrel 3 results in Figure 5.6 compared with Figure 5.5. Importantly, on comparing Figures 5.6 and 5.4, the improvement

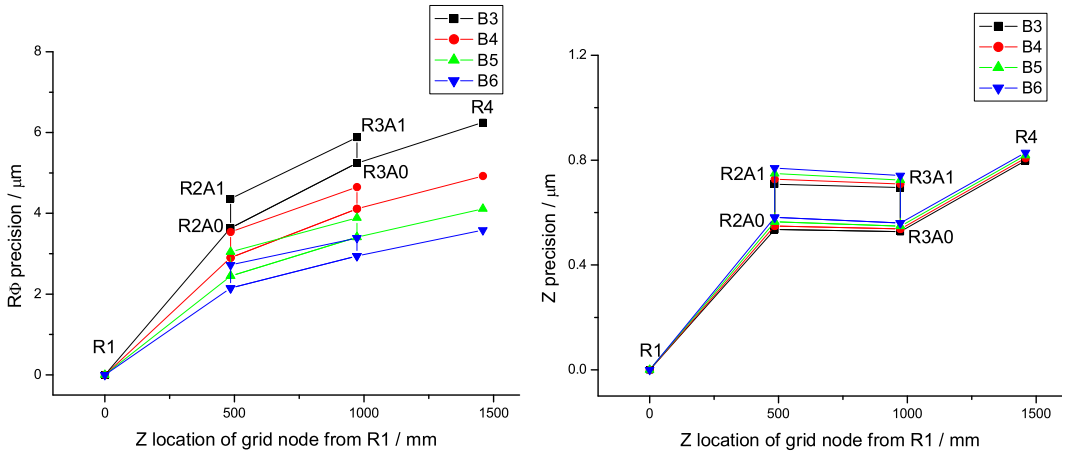

Figure 5.6: *The RΦ and Z precisions on the grid nodes of a quadrant grid for each Barrel 3-6. The rigidity of the barrel-flanges was modelled. The precision is plotted as a function of the grid node Z location from Ring 1.*

is greater for the quadrant than for the single octant. The extra redundancy of the additional adjacent octant improves the rigidity of the original octant.

The variation of RΦ and Z precisions with barrel radius is similar to that observed in the octant grid, for the same geometrical reasons.

Grid Section	Barrel-flanges	Number of GLIs	Model DoF	Total Redundancy	Redundancy per octant
Octant	non-rigid	11	9	2	2
	rigid	11	7	4	**4**
Quadrant	non-rigid	21	17	4	2
	rigid	21	11	10	**5**

Table 5.1: *Redundancy of the octant and quadrant models, with and without barrel-flange constraints.*

The quadrant simulations show that only when rigid flanges are modelled, does an adjacent octant create extra rigidity in the original octant. This can be explained by examining the DoF and redundancy in each model as in Table 5.1. When the flange-rigidity is added to the octant grid model, the redundancy increases by 2. When the flange-rigidity is added to the quadrant grid model, the redundancy *per octant* increases by 3, providing two *extra* redundant grid lines in the total quadrant. These are the source of the observed *extra* improvement in the rigidity of the overall quadrant and hence the original octant. The rigid flanges aid transmission of grid rigidity between octants.

(c) Inner Grid of Single Barrel

In the final step of this phase, the complete inner barrel grid was simulated for each barrel. A SIMULGEO model was written to describe all FSI components shown in Figure 5.1, comprising 4 rings of 8 FSI jewels. The SIMULGEO rendering of the model is shown in Figure 5.7. The GLIs of one quadrant of the barrel grid have been overlaid with red lines for clarity.

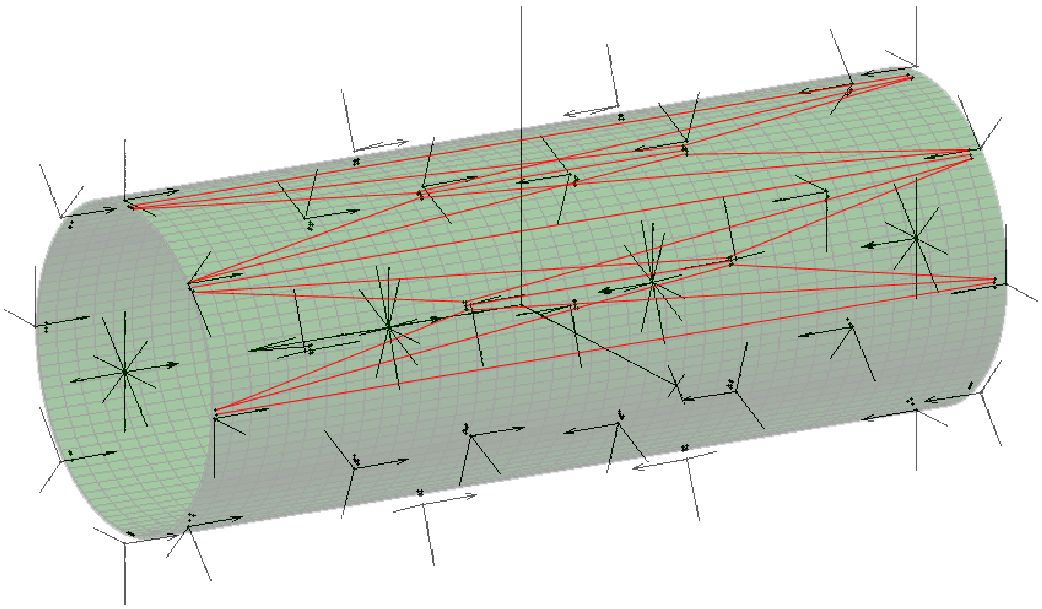

Figure 5.7: *A SIMULGEO model of the inner barrel grid, superimposed on an FEA model of Barrel 3. The grid line interferometers that monitor one barrel quadrant are shown.*

A multitude of ways to define the DoF in the model exist. A minimal set of simulations is presented here to highlight important results.

A simple model was defined in which the four rings of jewels were treated as distinct objects, each with 6 DoF, and no relative movement was allowed between the jewels within a ring. Ring 1 was the reference. Rings 2-4 were free in all 6 DoF. This simple model relates the movement of large structures and should enable a quick assessment of some basic eigenmode deformations of each barrel shape. The model has obvious limitations as it restricts the majority of translational DoF of the grid nodes.

The simulation was performed for all barrel radii. The transverse translational and axial rotational precisions on Rings 2-4 are compared for barrels of different radii in Figure 5.8.

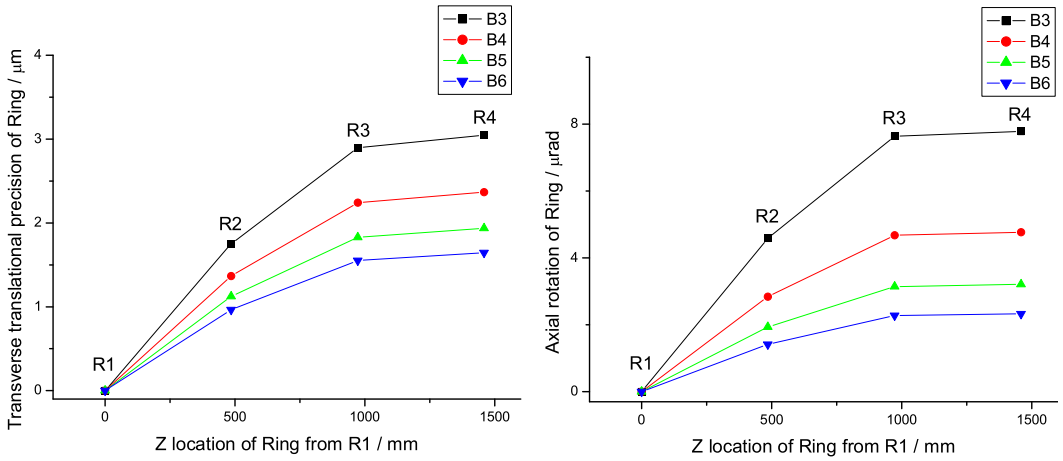

Figure 5.8: *Simulated precisions for a simple model of the inner barrel grid which assumed rigid rings of jewels. The left plot shows the translational precision in the RΦ plane of each Ring. The right plot shows the axial rotational precision on each Ring. The precisions are plotted versus the Z location of the Ring from the reference Ring, R1. The simulations for all barrel radii are compared.*

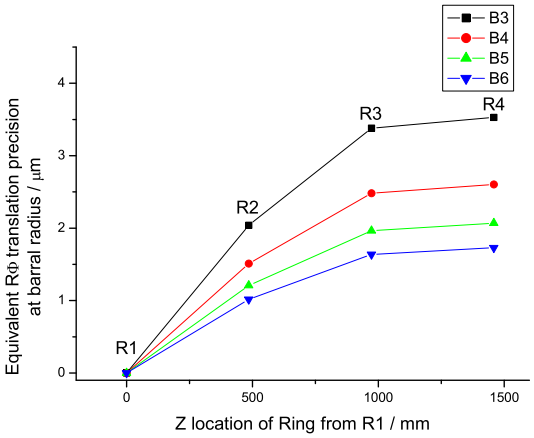

Figure 5.9: *The equivalent RΦ translational precision at the barrel surface, calculated from the combined precisions of Figure 5.8. The precision is plotted versus the Z location of the Ring from the reference Ring, R1. The simulations for all barrel radii are compared.*

Both types of precision worsen with the increasing separation in Z from the reference Ring, as explained by the tower block analogy described in the octant grid simulation phase. The precisions improve with barrel radius, due

to the reducing length:width aspect ratio of the octant grid. The equivalent translation at the surface of each barrel due to the axial rotation, is found by simply multiplying by the barrel radius. This was added in quadrature with the transverse translational precision for each ring to give the combined effective translation at the barrel surface, as in Figure 5.9.

The deformations of the real barrel will be far more complex than allowed for in this simple model. However, this model should be able to precisely determine the average position and orientation of rings of FSI jewels in a single barrel, for quick assessment purposes.

More complex deformations were allowed by defining a model of the inner barrel grid in which all jewels were constrained in R to 5 μm but free in RΦ and Z. The reference frame was formed by the 3 nodes R1A0, R1A4 and R4A0. Node R1A0 was rigidly fixed to the reference frame, R1A4 was rigidly fixed in Z and RΦ and R4A0 was rigidly fixed in RΦ. The three nodes therefore defined the XZ plane.

For this model, only results from Barrel 3 are presented, as this is the worst case, due to the aspect ratio of the barrel.

Figure 5.10 shows the RΦ and Z precisions on each jewel of a given ring, plotted against azimuthal position of the jewel. The node precisions worsen with azimuthal separation from the reference nodes. These degradation patterns have symmetry about the XZ plane, which is defined by the reference nodes. The RΦ and Z precisions for jewels in the central two rings are within the 10 μm required for ATLAS, even without the assumption of rigid barrel flanges. Extra redundancy gained by closing the grid loop around the barrel improves the grid rigidity.

The simulation was repeated omitting all four GLIs connecting the central jewels in octant A7 to the endjewels in octant A0, so that the redundancy loop around the barrel was cut completely. Figure 5.11 shows the impact of cutting the grid on the RΦ and Z precisions of each jewel as plotted against azimuthal position of the jewel.

A dramatic worsening in both precisions is observed with the azimuthal separation of the jewels from the reference nodes. With this definition of DoF, the precisions clearly fail to meet the ATLAS requirements on all grid nodes. The 7 remaining octants still individually meet the ATLAS requirement, but this simulation shows that the accumulative error of multiple adjacent octants is very large, when the rigidity of the barrel-flange is not assumed. The simulations show the importance of the closed redundancy loop of the grid around a single barrel.

The final model of the inner barrel grid presented here, includes the rigidity of the barrel-flanges. The jewels in the two central rings were constrained to 5 μm in R and free in RΦ and Z. Rings 1 and 4 were defined as extended

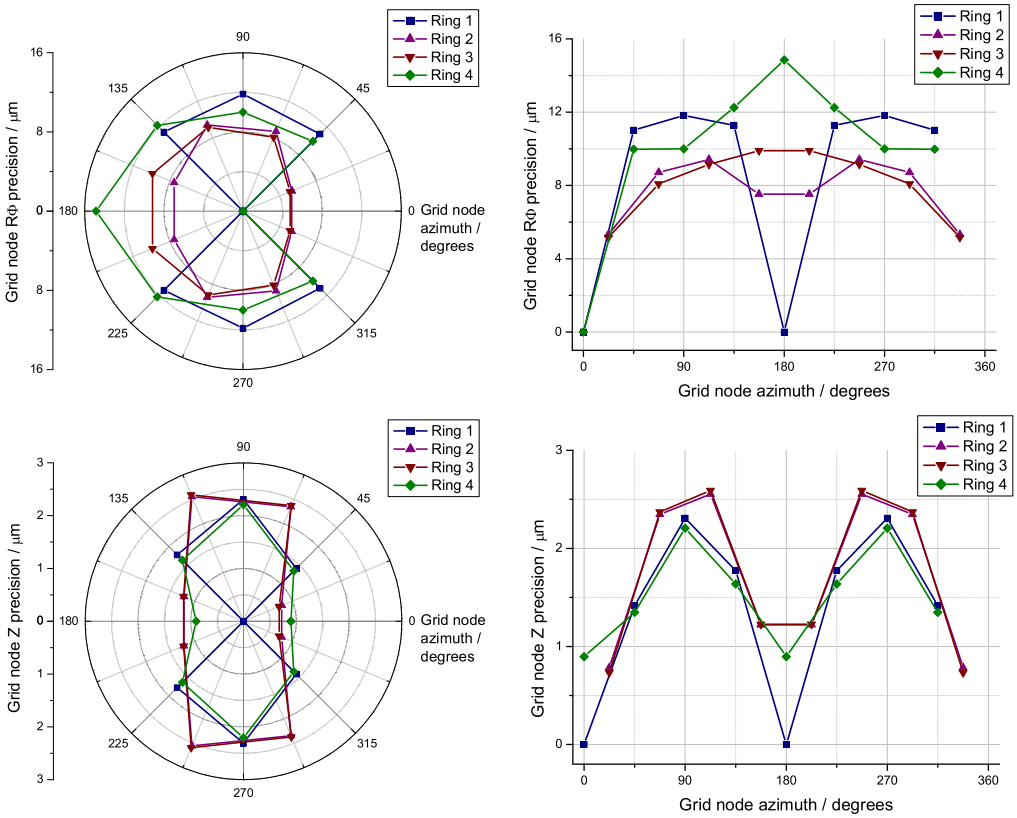

Figure 5.10: *Simulated precisions for the inner grid of Barrel 3, without assuming rigid barrel-flanges. The RΦ and Z precisions on each jewel of a given ring are plotted against azimuthal position of the jewel. Jewels R1A0, R1A4 and R4A0 were absolutely fixed in RΦ and jewels R1A0 and R1A4 were absolutely fixed in Z. The node precisions worsen between the reference nodes.*

objects, in which the jewels were all locally constrained to 2 μm in R, RΦ and Z, to model the barrel-flange rigidity [Lau01]. Ring 1 was the reference and Ring 4 was free in all 6 DoF.

The DoF definitions had rotational symmetry about the barrel Z axis, so the simulation yielded identical precisions on each jewel of Ring 2 or of Ring 3. The precision on each jewel in Ring 4 was calculated by summing in quadrature the relevant translational and rotational precisions on Ring 4 and accounting for each barrel radius. The equivalent translational precisions of each node are shown in Figure 5.12.

The simulated precisions are well within the ATLAS requirements, given the above assumptions on barrel-flange rigidity. The impact of the assumed constraints on the RΦ precision was investigated by repeating the Barrel 3

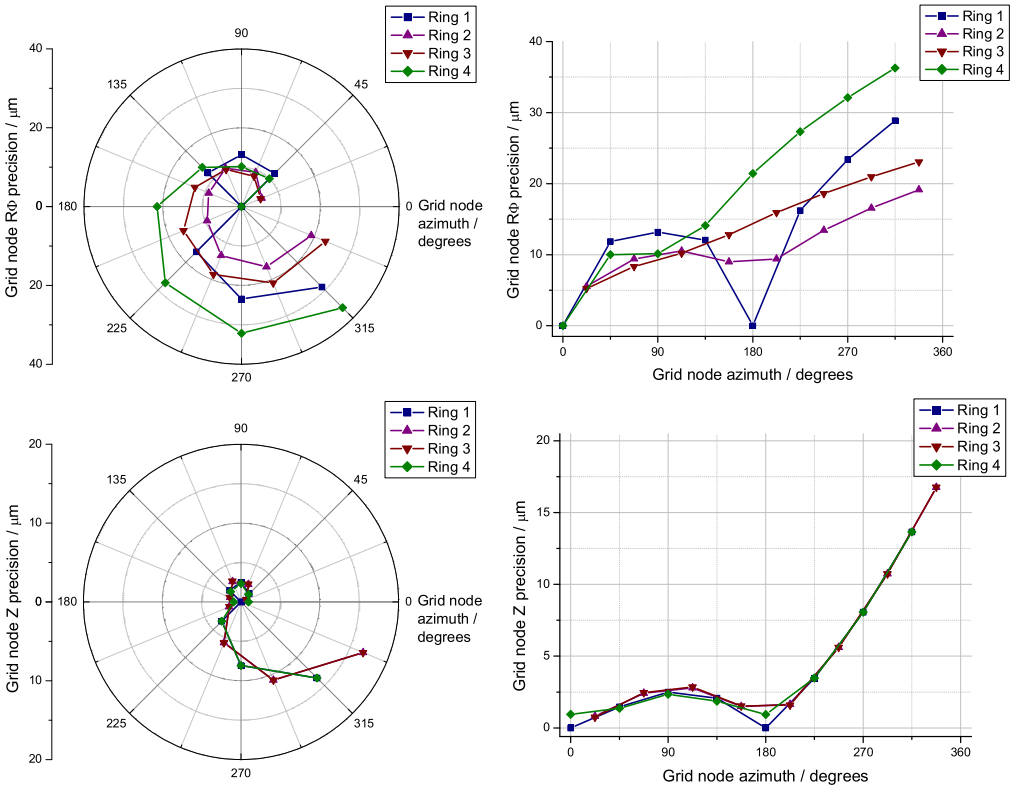

Figure 5.11: *Simulated precisions for the inner grid of Barrel 3, without assuming rigid barrel-flanges, in which the grid has been cut between octants A7 and A0. The RΦ and Z precisions on each jewel of a given ring are plotted against azimuthal position of the jewel. A dramatic worsening in both precisions is observed with the azimuthal separation of the jewels from the reference nodes.*

simulation and varying the value of the constraints in the model. The RΦ precision was found to not vary dramatically with the constraints in R or RΦ of each jewel, for Rings 1 and 4. However a significant variation was found on the RΦ precision of the nodes in the each central ring with the Z constraint on the endjewels in Ring 1 and 4, as in Figure 5.13. This variation indicates the reliance of the barrel grid reconstruction on the rigidity of the barrel flange in Z.

The inner barrel grid simulations indicate that the ATLAS requirements should be met if the described assumptions on the barrel-flange rigidity and the local deformations of the carbon-fibre cylinders prove to be true. If the assumptions are flawed, models with fewer assumptions still indicate valuable information may be obtained from the FSI alignment system. The simulations have also revealed some weaknesses and strengths of the inner barrel grid.

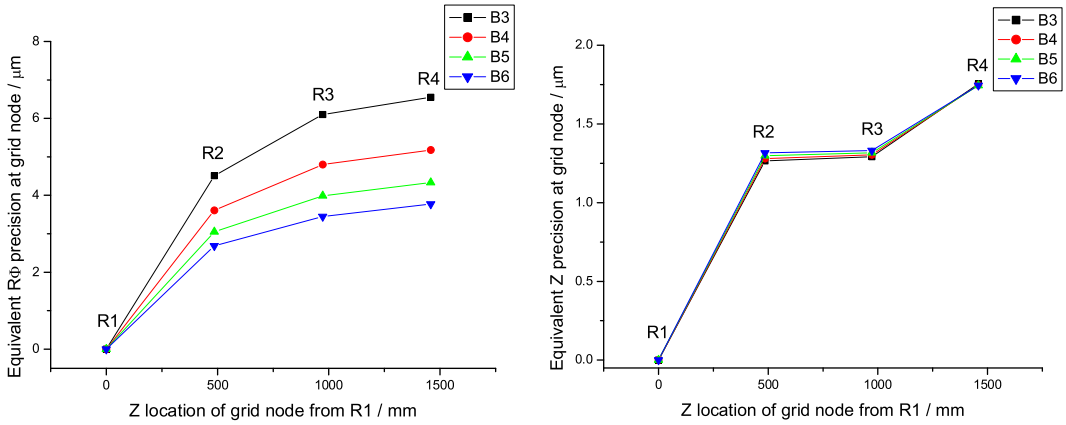

Figure 5.12: *The equivalent RΦ and Z precisions of the jewels in Rings 1-4, versus the Z location of the jewel from Ring 1. The precisions for all barrel radii are compared. An endjewel Z constraint of 2 μm was assumed.*

Figure 5.13: *The variation in RΦ precision on the nodes in Rings 2 and 3 with the Z constraint on the endjewels in Ring 1 and 4. The simulation was for Barrel 3.*

5.3.4 SCT Barrel-Flange Grid

The SCT barrel-flange grid, shown in Figure 5.14, consists of 72 FSI measurements that connect the four barrels in the transverse plane at one end of the SCT barrel.

The primary purpose of the barrel-flange grid is to monitor the relative positions and orientations of the four barrels. These DoF are particularly difficult to reliably determine with tracks, because progressive relative translations and/or rotations of the four barrels lead to sagitta deformations, discussed in

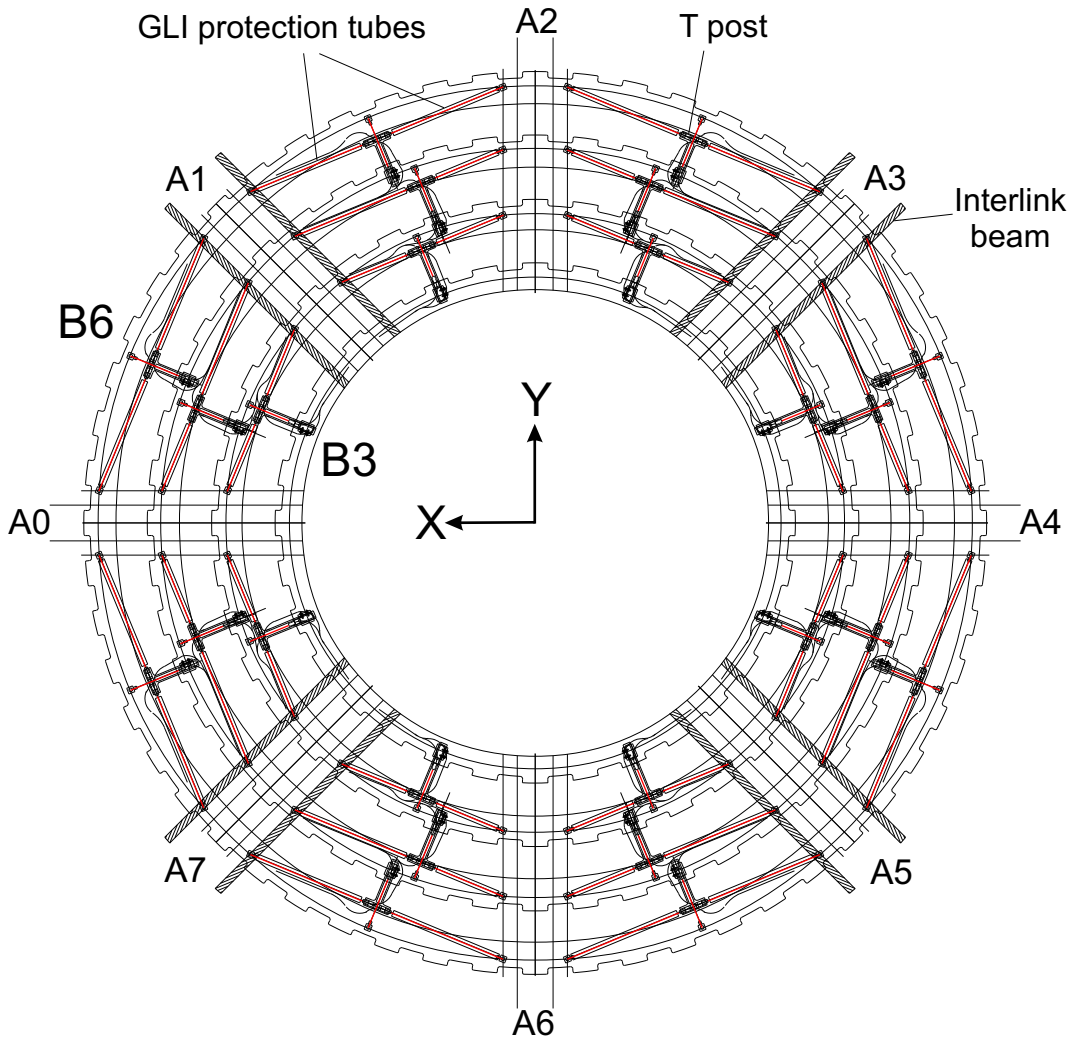

Figure 5.14: *The ATLAS SCT barrel-flange alignment grid at one end of the four barrels. The quills are held in T-shaped posts that are attached to barrels 3-5. Each quill forms a GLI with a retro-reflector on the adjacent barrel. The 72 GLIs are represented by red lines. The GLIs run inside tubes that protect the lines of sight against obstruction.*

Section 5.2.1. It is therefore essential that the FSI alignment system monitors these DoF with excellent precision.

The secondary purpose of the barrel-flange grid is to help constrain the shape deformations of each barrel-flange, to aid the internal alignment of each barrel.

The barrel-flange grid contains 8 pairs of retro-reflectors per barrel (except for barrel 3) that straddle the 8 interlink beams. Each retro-reflector pair is

located near to one of the 8 endjewels of the inner barrel grid, on the opposite side of the flange. Eight further retro-reflectors per barrel, excluding barrel 3, are located on the flange about midway between the interlink beams. The quills are held in T-shaped posts that are attached to barrels 3-5. Each quill forms a GLI with a retro-reflector on the adjacent barrel. The T-posts are not located near to any node of the inner barrel grid.

The barrel-flange grid was modelled in SIMULGEO. Each barrel-flange was defined within a distinct co-ordinate frame, that was allowed to translate and rotate only in the RΦ plane. All flanges were rigidly fixed in Z and in rotation about X and Y axes, effectively creating a two-dimensional grid.

The T-posts were defined in distinct co-ordinate frames that were fixed to the relevant flange with a 2 μm constraint in R, RΦ and Z, to model the barrel-flange rigidity [Lau01]. The orientation of a T-post with respect to its flange was fixed absolutely. The retro-reflector pairs that straddle the interlinks had identical constraints to the T-posts. No relative movement was allowed between quills within a T-post or between the retro-reflectors within a pair that straddles an interlink. The remaining retro-reflectors were individually fixed to the relevant flange by 2 μm in all 3 DoF.

In repeated simulations, each barrel-flange was systematically defined as the reference. The model therefore solved only 9 free DoF, corresponding to the 2 translations and 1 rotation in the RΦ plane of each other flange. The 2 translational precisions of a flange were identical, due to symmetry. The simulated precisions are plotted in Figure 5.15.

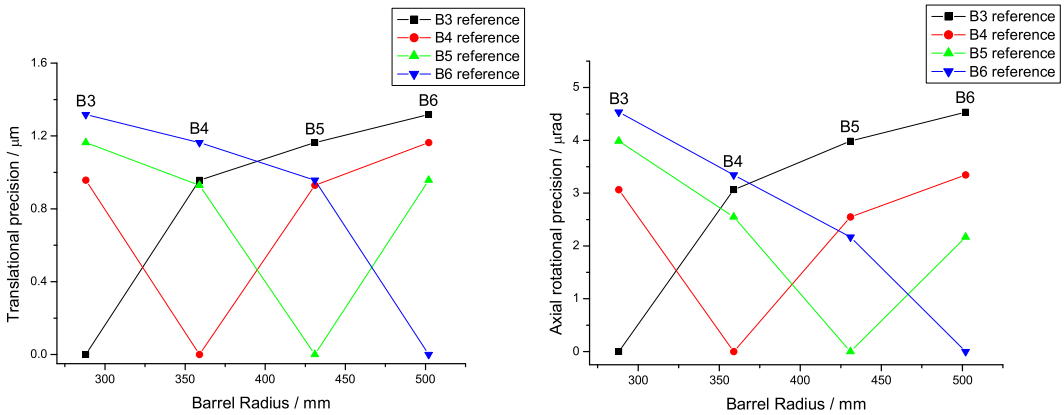

Figure 5.15: *The translational and rotational precisions on each barrel-flange in the RΦ plane, for a simulation of the SCT barrel-flange grid that modelled the rigidity of the flanges. The simulation was systematically repeated with each barrel flange defined as the reference, while the other barrel flanges were free.*

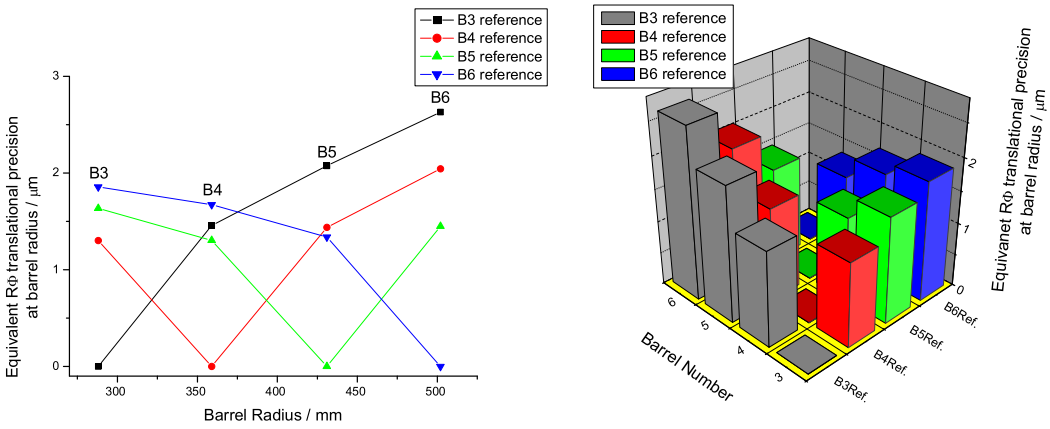

Figure 5.16: *The equivalent translational precision at the barrel surface, calculated from the combined precisions of Figure 5.15, for a simulation of the SCT barrel-flange grid that modelled the rigidity of the flanges.*

The simulated precisions are well within the ATLAS requirements. The precisions degrade with increasing radial difference between the flange and the reference flange. The pattern of degradation for the translational precision of each flange, shown in Figure 5.15, is symmetric; for example, the translational precision of B5 with respect to B3 is the same as B4 with respect to B6. The corresponding symmetry is skewed for the axial rotational precision; for example the axial rotation of B4 with respect to B6 is more precisely determined than for B5 with respect to B3.

An estimate on the translational RΦ precision at the barrel surface was calculated by summing in quadrature the translational and rotational precisions on each flange, after accounting for the relevant barrel radius. This equivalent translational precision has skewed symmetry as shown in Figure 5.16. The skewed symmetry is due to the differing barrel radii and lengths of the T-posts on each barrel.

The simulated precisions were found to be very sensitive to the barrel-flange rigidity constraints. When the flanges were modelled to be completely rigid, all simulated precisions improved on the above by a factor of 3 and retained the patterns of degradation. The rigidity of the real ATLAS barrel-flange is therefore critical to the precision achievable with this model.

Limitations and future work

The above model should allow a quick assessment of the average position and orientation of the barrel-flanges. However, the model only solves for 9 DoF with

72 GLIs, so clearly neglects the majority of information that the grid provides regarding the shape deformation of each flange. More DoF of the grid nodes need to be released to allow the shape deformations in the model. However, difficulty arises in defining appropriate DoF with SIMULGEO, due to the grid configuration and the limitations of the software.

A SIMULGEO model that would allow some deformations of each barrel-flange could be defined by releasing the R and RΦ co-ordinates of each pair of retro-reflectors that define a node on an interlink. The small R component of the GLIs from the T-posts to the interlink nodes, should allow the R co-ordinate to be determined, in addition to the RΦ co-ordinate of the interlink nodes. This should help constrain the endjewels on the opposite side of the flange to the interlink nodes. However, no such measurements exist between the interlinks on Barrel 3.

Within each sector of the barrel-flange grid, the radial GLIs emanating from the 3 T-posts connect 4 nodes that lie on an approximately radial line between the interlinks. However, the radial co-ordinates of these nodes are not related to the rest of the grid by any direct FSI measurements. These nodes are also not located near to any other grid nodes, for example of the inner barrel grid, that might constrain them. A completely free model of the barrel-flange grid, therefore does not solve.

However, due to the carbon-fibre barrel-flange, the radial position of the 4 nodes between the interlinks should be correlated with the radial positions of the nodes on the interlinks. The exact correlation is likely to be complicated and dependent on the eigenmode deformations of the SCT barrel. In a naïve model, the radial position of each node could be locked to the average radial position of the two adjacent interlink nodes. However, SIMULGEO does not permit the average position between grid nodes to be defined, particularly due to further complications created by the closed loop of nodes around the barrel.

A possible solution is to fit circular eigenmodes directly with the FSI measurements to determine the shape deformations. The effect of local deformations of the T-post, should also be investigated.

Additionally, FEA models indicate that the positions of the 4 nodes between the interlinks are not correlated with the positions of just the interlink nodes, but all nodes in the SCT barrel grid, due to eigenmode deformations of the complete structure [Cug01]. In the future, it is planned to extend the simulations of the barrel-flange grid to incorporate knowledge of the structure gained through measurements and FEA models.

5.3.5 Complete SCT Barrel Grid

The progressively complicated simulations of the previous sections culminated in the simulation of the complete SCT Barrel grid, consisting of 4 inner barrel grids and 2 barrel-flange grids. The SIMULGEO rendering of the model is shown in Figure 5.17. The GLIs of one quadrant of the inner grid of barrel 6 and one quadrant of one SCT barrel-flange grid have been overlaid with red lines for clarity.

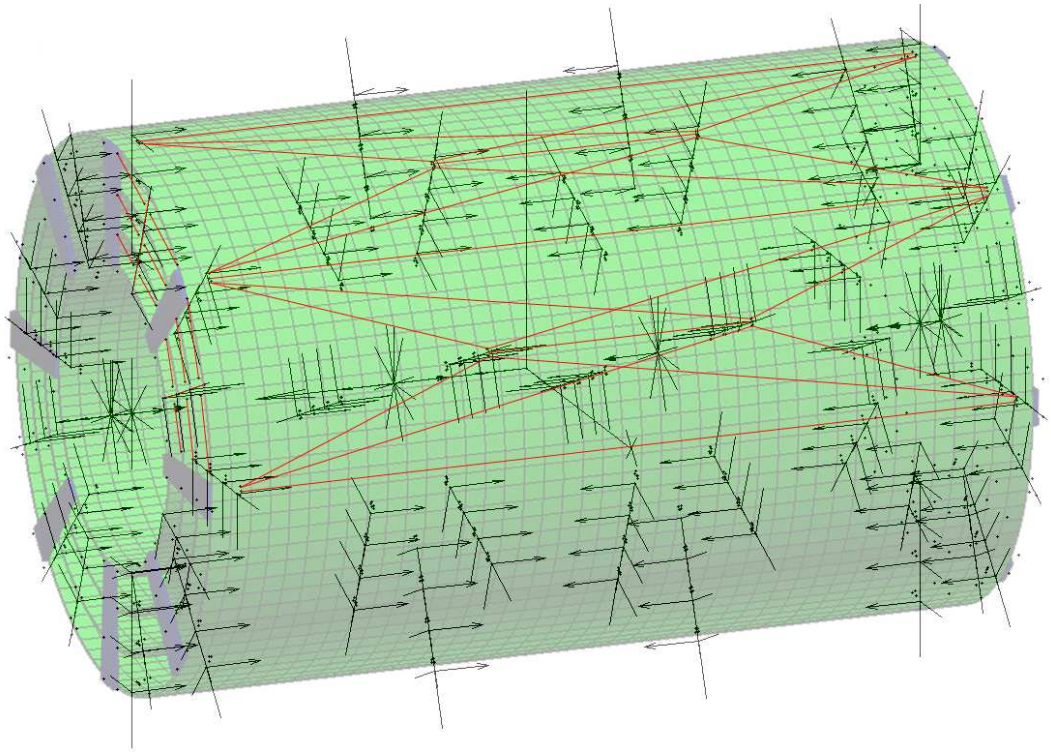

Figure 5.17: *A SIMULGEO model of the complete SCT barrel grid, superimposed on an FEA model of the SCT barrel. The red lines represent the 39 grid line interferometers that monitor one quadrant of the outer barrel and one quadrant of one SCT barrel-flange. The total SCT barrel grid comprises 512 GLIs.*

Assumptions of the model

The definitions in the complete SCT barrel grid model combined aspects of the most advanced models of each previous simulation phase. The four barrels were defined in distinct co-ordinate frames. A barrel comprised 6 rings of jewels, corresponding to the four rings of jewels of the inner barrel grid and

two rings of jewels for the barrel-flanges. Rings 1, 4 and the two barrel flange rings were defined as distinct frames within each barrel. However, Ring 1 and the adjacent barrel-flange ring were rigidly locked together, as were Ring 4 and its adjacent barrel-flange. Thus each ring of endjewels and adjacent barrel-flange ring created one three-dimensional composite object, with 6 external DoF. These objects are referred to as Ring 1 and Ring 4 in the rest of this section.

Each endjewel in Rings 1 and 4 was rigidly fixed to the pair of retro-reflectors nearest to that endjewel on the opposite side of the flange, to create one grid node. This node was locally constrained to its ring, to 2 μm in R, RΦ and Z, to model the expected barrel-flange rigidity [Lau01]. The T-posts and other retro-reflectors of the barrel-flange grid were also locally constrained to 2 μm in R, RΦ and Z.

The jewels in the central two rings, R2 and R3, were free in RΦ and Z and constrained to 5 μm in R. In the SCT barrel grid design there are GLIs that measure from each central jewel to a retro-reflector on the nearest module, radially inward from the jewel. These GLIs were omitted from the model described here, so that the results from the SCT barrel grid could be directly compared with the results from the inner barrel grid simulation, which also omitted these GLIs.

No relative displacement was allowed between the measurement nodes within any jewel. All jewels were locally constrained almost completely (to 0.1μrad) in their orientation.

All precisions were calculated with respect to Ring 1 of Barrel 3, which was defined as the reference. Each Ring 4 of all barrels was free in all 6 DoF. Each Ring 1 of Barrels 4-6 was free in 5 DoF, but the Z coordinate was constrained almost completely (to 0.1 μm) to the Z location of Ring 1 of Barrel 3. The Z constraint is discussed below.

The grid was solved assuming a 1 μm measurement precision on all GLIs. Around 5 minutes were required to read the model into SIMULGEO and the calculation took less than 30 seconds, on a 1.6 GHz Pentium® 4 processor. The calculation consumed 7,706 kbytes of memory. The calculation time is dominated by the inversion and diagonalization of the covariance matrix [Bru98]. The covariance matrix rank was 922, due to solving all free and constrained parameters, for the best estimate on the precisions. When only the free parameters were solved and the constrained parameters were assumed absolutely rigid, the covariance matrix rank was only 167 and the calculation time was less than 1 second.

Simulation results and discussion

The grid simulation yielded identical precision estimates for each jewel within Ring 2 or within Ring 3, as expected by symmetry. For Ring 1 or Ring 4, the simulation yielded identical translational precisions in X and Y, and identical rotational precisions about X and about Y, also expected by symmetry. The translational RΦ and Z precisions on a grid node at the Barrel surface was estimated for each barrel by summing in quadrature the translational and rotational precisions on Ring 1 or Ring 4, after accounting for the relevant barrel radius. These equivalent RΦ and Z translational precisions for Rings 1 and 4, and the individual RΦ and Z translational precisions on any grid node of Rings 2 and 3, are plotted for each barrel in Figure 5.18.

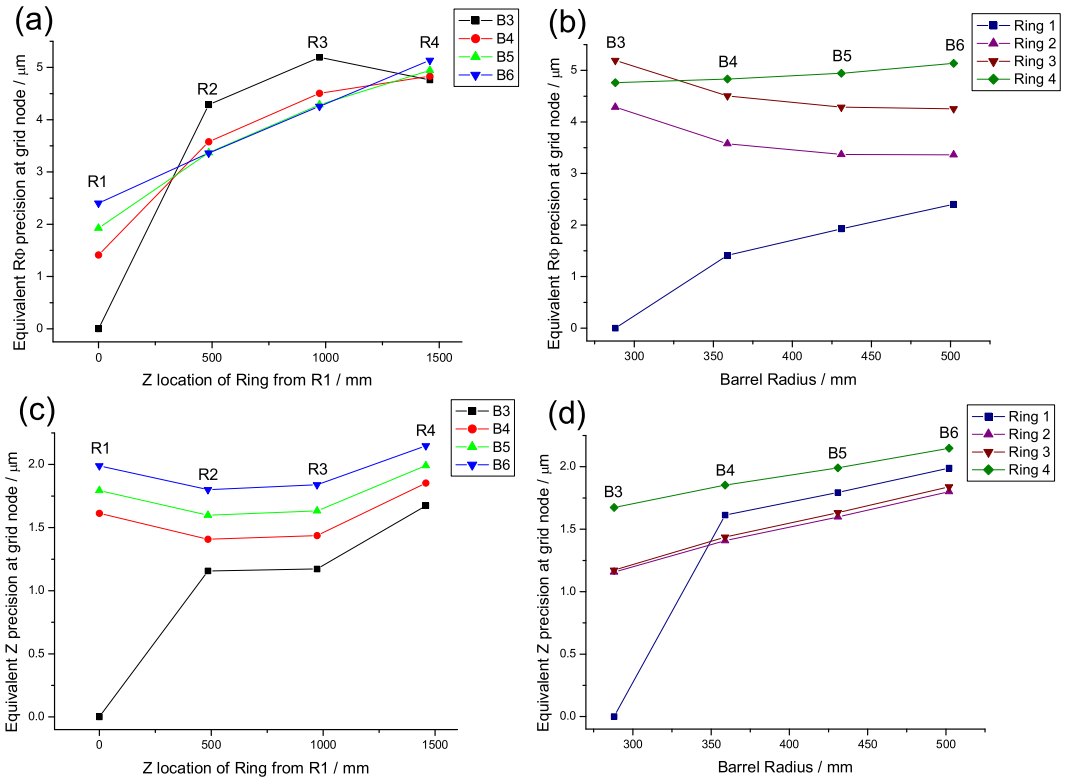

Figure 5.18: *Simulated precisions for the complete SCT barrel grid, in which the rigidity of the barrel flanges was modelled. Ring 1 of Barrel 3 was the reference. The equivalent RΦ and Z translational precisions on a grid node from each Barrel and Ring are plotted on the left versus the Ring Z location from the reference Ring and versus the Barrel radius on the right.*

Notably, all precisions are within, or very close to within the most stringent

5 μm RΦ ATLAS requirement. The Z precisions are about 2 μm or better.

The precisions for the complete SCT barrel grid were sequentially examined and compared with the relevant results from the previous simulation phases.

When the RΦ precision for the SCT barrel grid, shown in Figure 5.18(a), is compared to that of an isolated Barrel 3 inner grid, shown in the left plot of Figure 5.12, an improvement on the RΦ precision is apparent for Ring 4 of Barrel 3. The RΦ precisions on Rings 2 and 3 for Barrel 3 are also improved slightly. The precision has improved because the extra constraints of multiple rigid barrel-flanges increase the extra redundancy of the SCT barrel grid compared with inner barrel grid. The effect is comparable to the effect of extra redundancy previously noted between the isolated barrel octant and quadrant grids. For the SCT barrel grid, the additional redundancy of the adjacent barrels clearly improves the rigidity of each barrel grid.

Interestingly, for Barrel 3 in the SCT barrel grid, the equivalent RΦ precision on Ring 4 is better than on Ring 3, despite Ring 4 being further from the reference Ring. Extra redundancy causes the complete Ring 4 of Barrel 3 to be very well determined: more so than for the RΦ precision on the individual nodes of Ring 3 of Barrel 3, due to the poor aspect ratio of the octant grid of Barrel 3. The octant aspect ratio is better for Barrels 4-6, so the RΦ precision on Barrels 4-6 degrades with Z displacement from the reference Ring, as expected.

Figure 5.18(b) shows the RΦ precision of a node in each Ring plotted against the Barrel radius. When these data are compared with the B3 data in the left plot of Figure 5.16, a small improvement on the equivalent RΦ precision is apparent on Ring 1 of Barrels 4-6. This is again due to increased rigidity from the extra redundancy of multiple rigid barrel-flanges. The precisions on Rings 2 and 3 reduce with barrel radius, due to the reducing length:width aspect ratio of the octant grid.

The equivalent Z precisions on nodes from each Ring in Barrel 3 of the SCT barrel grid, in Figure 5.18(c), are almost identical to those of the isolated Barrel 3 inner grid, shown in the right plot of Figure 5.12. The Z precisions do not improve because there are no extra Z constraints between barrels.

Figure 5.18(d) shows that Z precisions on each Ring of the Barrels, degrades approximately linearly with Barrel radius.

Impact of constraints

The above model constrained almost completely (to 0.1 μm) the Z co-ordinate of each Ring 1 of Barrels 4-6, to that of Ring 1 of Barrel 3. The simulation was repeated with this constraint loosened to 10 μm. The effect was to degrade all Z precision on Barrels 4-6 by around 10 μm (a quadrature sum of 10 μm

with the originally simulated values), but the RΦ precisions were negligibly affected. The grid did not solve when these Z DoF were completely released because there are no GLIs in this model that measure between barrels with any Z component. However, it is expected that these Z constraints should not be necessary when the omitted GLIs, which have a significant Z component, are included.

Further simulations investigated the effect of releasing the angular constraint about Z, of all jewels in Rings 2 and 3. This constraint was progressively loosened from 0.1 μrad in orders of magnitude. A constraint of 100 μrad or less had very little effect on any precision. A constraint of 1000 μrad, degraded the RΦ precision on all jewels in Ring 2 to 13.0-14.6 μm and in Ring 3 to 13.4-15.4 μm. All Z precisions were very little affected. FEA models indicate the maximum rotation of the central jewels about this axis should be approximately 100 μrad [Cug01].

Summary of SCT Barrel Grid Simulations

Simulations of the near complete SCT barrel grid indicate that the required precision for the ATLAS FSI alignment system should be achieved, if certain assumptions on the rigidity of the SCT support structure prove to be true. In the future, it is planned to develop these models, to reduce the need for some of these assumptions or to confirm them with measurements of prototype support structures.

5.4 SCT End-cap Grid Simulations

5.4.1 Overview

This section summarizes how the SCT End-cap alignment grid was designed and configured to fulfill its role in ATLAS. The layout of the SCT End-cap and the ATLAS requirements are first reviewed, the design history is summarized and the ATLAS SCT End-cap design that is currently in production is evaluated.

5.4.2 End-cap Layout and Alignment Requirements

Both SCT End-caps are identical in design and should have rotational symmetry about an almost vertical line through the interaction point when installed in ATLAS. Each SCT End-cap consists of nine wheels of modules that are coaxial with the beam-line. The modules of a wheel are secured to a carbon-fibre disc in three annular rings of increasing radius. The modules of an inner and outer ring are mounted on the opposite side of the disc to the middle ring of modules, ensuring complete coverage. The separation in Z between discs is configured to maximize the efficiency of at least four SCT silicon plane transitions per track [Hay02].

The nine wheels of modules are each mounted via 12 special disc fixings [Wer03] within a support cylinder. The cylinder is suspended via two wings at four points on the squirrel cage rails, attached to the solenoidal magnet. Small movements of the squirrel cage rails would distort the SCT End-cap and induce relative shifts and internal deformations of the module wheels. The FSI system should monitor these distortions and others induced by varying thermal gradients, dehumidification and cooling pipe forces, with high precision.

The alignment requirements for the SCT End-cap are the same as those discussed in Section 5.2.1, except that the Z and R requirements are modified to 200 μm and 50 μm respectively [ATL97].

The layout imposes awkward constraints on the FSI alignment system. There is almost no free space in which to mount FSI components on the side of the disc that has two rings of modules. The largest FSI components are therefore installed on the side of the disc with only one ring of modules.

5.4.3 Grid Design History

The original SCT End-cap FSI grid design [Huf99] was abandoned, owing to the design changes caused by the need for an insertable pixel detector. The design

radius of the inner thermal enclosure increased, obstructing critical lines of sight in the grid. The grid was completely redesigned, based on grid simulations performed by the author.

Several very different grid configurations were envisaged and compared. The primary requirement was for lines of sight that connected at least three well separated nodes on a disc to one node on the adjacent disc, to allow three dimensional reconstruction of the node on the adjacent disc. Designs were considered and excluded on the following grounds:

- A grid with four nodes per disc and a square cross-section, is not sufficiently rigid (diagonal cross-bracing is obstructed by the thermal enclosure).
- A grid with four nodes per disc at the corners of a square, rotated by 45° on alternate discs allows all possible connections of nodes on adjacent discs for increased rigidity. However, several lines of sight are obstructed by the inner thermal enclosure.
- A pentagonal design, with an equivalent rotation of nodes on alternate discs, does fit around the inner thermal enclosure, but significantly increases the number of GLIs required.

The initial studies converged on a grid design with an equilateral triangular cross-section, in which 6 outer nodes on adjacent discs are connected by all $\frac{n(n-1)}{2} = 15$ possible lines of sight. Such a grid is therefore rigid and redundant. The lines of sight have just enough clearance around the inner thermal enclosure and are efficiently few in number, given the rigidity and redundancy this configuration provides.

5.4.4 Grid Connecting Adjacent Discs

For the grid with a triangular cross-section, the arrangement of GLIs between adjacent discs is shown in Figure 5.19. This forms the baseline design of the final grid.

The three outer nodes of a disc are sufficient to define a co-ordinate frame for the disc that has 6 external and 3 internal DoF. The outer nodes of adjacent discs are connected by 9 GLIs, shown by red lines in Figure 5.19. The 9 measured lengths suffice to determine the 6 external and 3 internal DoF of the adjacent disc, if the relative positions of nodes in the first disc are known. Three GLIs connect the outer nodes to form a triangle in the plane of the disc, shown by green lines in Figure 5.19. These in-plane triangles add redundancy and help determine the 3 internal DoF of the outer nodes of the disc. The

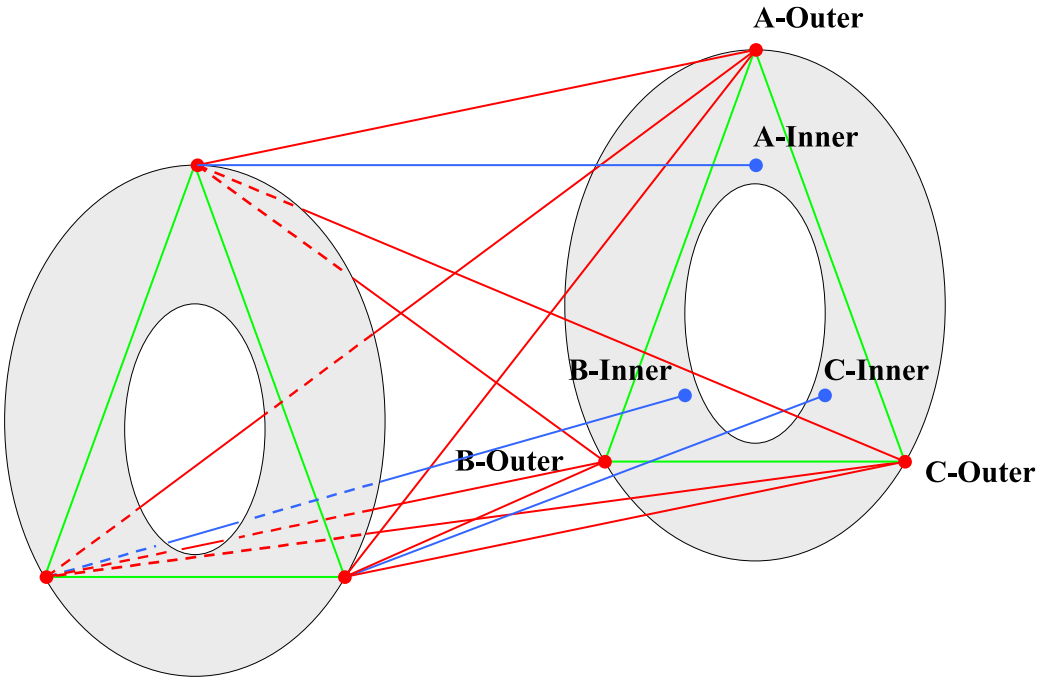

Figure 5.19: *The basic FSI grid configuration between adjacent discs of the SCT End-cap.*

15 GLIs between the outer nodes thus form a rigid geodetic grid that is repeated between all nine discs of the SCT End-cap. This rigid triangular-prism cage overdetermines the relative positions and orientations of the discs and the relative positions of the 3 outer nodes within the plane of each disc.

The discs are held in tension at twelve points around the outer rim of the disc. FEA models have shown that the low frequency eigenmodes (the lowest eigenmode frequency was calculated as 22.1 Hz) of the disc cause the inner rim of the disc to deform predominantly out-of-plane in the Z direction [Snip03]. The grid was therefore initially configured with additional measurements to 3 nodes on the inner rim of the disc, as shown by blue lines in Figure 5.19, to monitor these out-of-plane disc deformations.

The baseline grid was evaluated with a SIMULGEO model. The geodetic grid between the outer nodes was found to provide excellent rigidity. The simulations indicated that lateral displacements of the outer nodes on the last disc could be monitored with respect to the first disc, with a precision of less than 7 μm and the displacements in Z to less than 2 μm. More detailed results are documented elsewhere [Gib01b].

Notably, the proximity in Z between the two discs nearest the ATLAS inter-

action point (discs 1 and 2) significantly degrades the precision in determining the out-of-plane deformation of these discs. The precision on the local Z co-ordinate of an inner node of a disc varies with Z separation of the outer and inner nodes of adjacent discs, as in Figure 5.20, due to the varying Z component of the GLI between these nodes. The R-component of this GLI also varies, so any radial movement of the inner node, which was assumed constant in the model, would induce a systematic error in determining the Z co-ordinate of the inner node. This error would increase with the proximity of the discs, due to the increasing R:Z component ratio of the GLI.

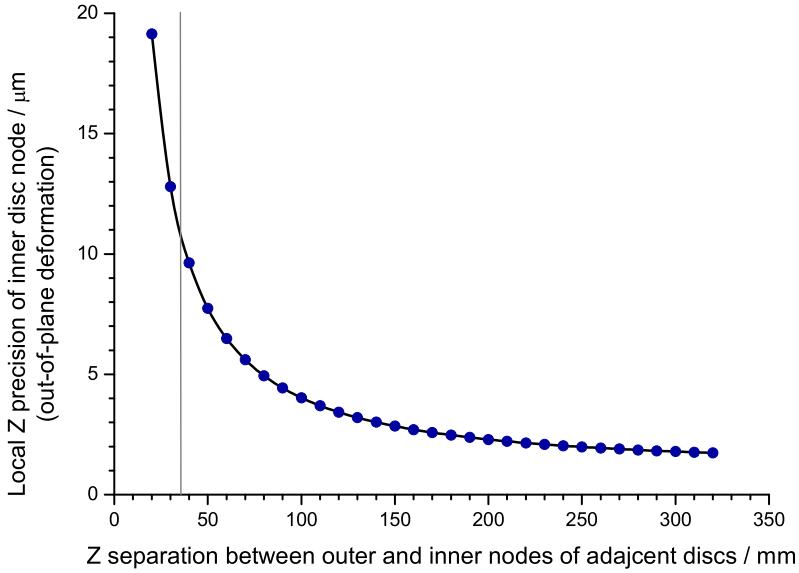

Figure 5.20: *Simulated precision on the out-of-plane deformation of a disc, for a preliminary design of the end-cap grid. After accounting for the dimensions of the FSI jewels in the preliminary design, the relevant node Z separation for discs 1 and 2 was 35.5 mm, as indicated by the vertical line, leading to a 10.8 µm precision.*

5.4.5 Design Evolution

Further grid designs were therefore created and studied to improve the precision described above. The further studies were also stimulated by the following reasons:

- the proximity of the first two discs also degraded the precision on outer nodes.

- the relative separation of particular measurement nodes within a jewel slightly degraded some precisions, so an improved design was sought.

- technical difficulties arose in constructing small enough FSI jewels to fit in the limited space in Z and maintain both the minimum optical-fibre bend radius and a sufficient length for the short GLIs to be well determined with FSI.

- the length of optical fibre for the FSI system was specified to ensure enough fibre for the baseline grid design was manufactured within tolerances. When the fibre was purchased, the fibre exceeded the baseline length and allowed some additional GLI channels to be included.

Instructive grid design failures

Various designs were considered and simulated, but shall not all be discussed here. However, particular design failures aided understanding of the strengths and weaknesses of the final grid design. One such interesting design was formed by inverting the layout of outer nodes between proximate pairs of discs (1 & 2 and 4 & 5), in an attempt to avoid the problem of short GLIs in Z between discs. In the model, the translational precisions on each disc were calculated with respect to a reference disc. The disc defined as the reference was systematically varied in repeated simulations and the calculated translational precisions are plotted in Figure 5.21.

A clear weakening of the grid was observed between disc pairs where the inversion of the node layout was applied. When disc 1 was the reference the translational precision of all other discs dramatically worsened. When disc 2, 3 or 4 was the reference, the dramatic worsening only started at disc 5 (disc 1 is also relatively poor). When disc 5 or higher was the reference, the dramatic worsening started at disc 4 and progressed in the opposite Z direction. These observations revealed there were poor linkages between discs 1 & 2 and 4 & 5, that lead to a large angular error about X and about Y between discs of an adjacent pair. The discs on the far side of these poor links were therefore increasingly free to sway in the transverse plane with respect to the reference disc, due to hinging about the weak connections.

These results indicate that inversion of the outer node layout between adjacent discs induces poor grid rigidity. In addition, the failed design demonstrates that the reference frame definition in the model can significantly alter the apparent precisions on the grid nodes. However, if all models have identical constraints, the grid shapes reconstructed with any of the models for a set of GLI lengths, should be identical and related simply by co-ordinate transformations.

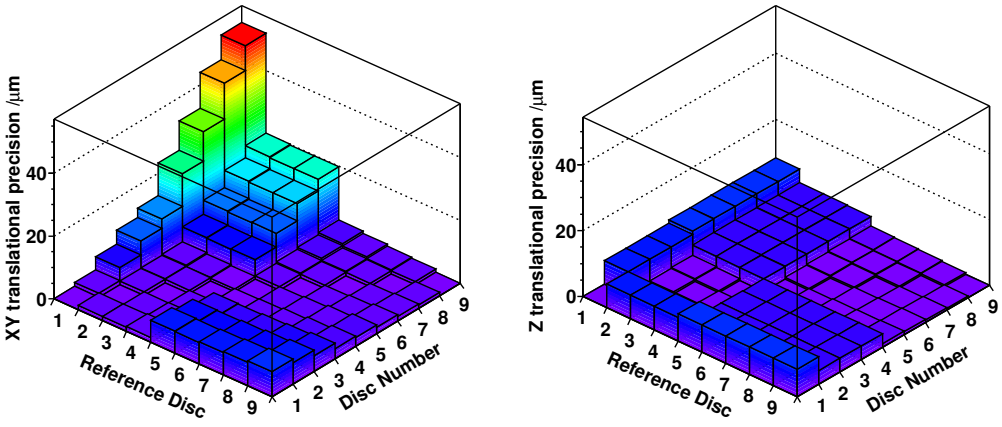

Figure 5.21: *Precisions on the translational precisions on the SCT end-cap discs, as simulated with a poor end-cap grid design. The precisions are plotted versus each disc, for nine simulations in which each disc was systematically defined as the reference. Identical axis scales allow a clear comparison.*

Additional grid nodes

In the baseline design described in Section 5.4.4, the three nodes at the inner radius, were connected to the well determined outer nodes by only one GLI, to monitor the out-of-plane deformations of the disc. This configuration suffers from the proximity problems discussed in Section 5.4.4. In addition, the measurements of out-of-plane deformation provide little data on the in-plane deformation of modules within a wheel, particularly in the RΦ direction, which is required to fine precision.

However, by including three additional nodes on the inner rim of a disc, as shown by the nodes labelled type-B in Figure 5.22, two GLIs may be formed between these nodes and the well determined outer nodes of the adjacent disc. These GLIs measure both the local Z and RΦ co-ordinates of the type-B nodes within a disc, to simultaneously monitor the out-of-plane and in-plane disc deformations. These measurements are also orthogonal to the radial component of the type-B nodes, so should not be degraded by small radial movements of the inner rim of the disc, as might be the case for the type-A nodes.

The very different disc separations in Z and the variation in number of modules rings on the last three discs, complicated the design process and layout of the final design. Several iterations of grid design were required to optimize the configuration of GLIs from the outer nodes to the various inner nodes. The following constraints were adhered to:

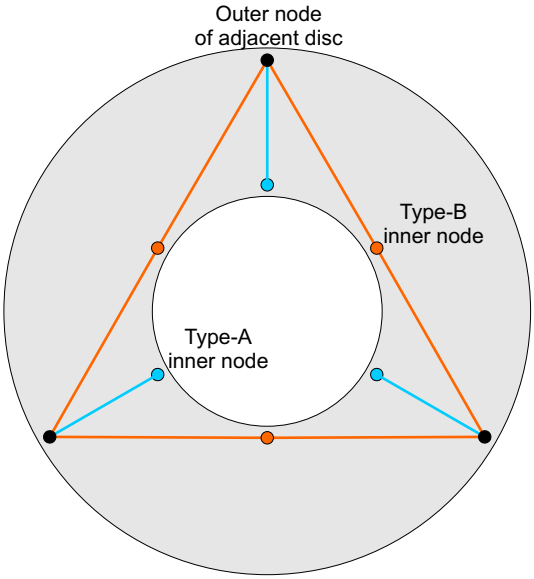

Figure 5.22: *The possible GLIs between the 3 outer nodes and 6 inner nodes of adjacent SCT End-cap discs. The two GLIs connected to Type-B nodes monitor the Z and RΦ movement of the node. Connections are also possible between directly opposite inner nodes of adjacent discs.*

- the allocated envelopes for FSI components.
- a sufficient clearance for all lines of sight.
- suppression of potential optical cross-talk between GLIs.
- an integer number of fibre ribbons per disc with minimal redundant fibre.
- limited total GLI channels due to FSI electronics constraints.

Connections directly between the inner nodes on adjacent discs were also optimized for redundancy. The iterations converged on the final design as presented and evaluated in the next section.

5.4.6 SCT End-cap Grid Evaluation

Grid layout

The alignment grid for each SCT End-cap contains 165 length measurement lines, as shown in Figure 5.23. The grid layout has a triangular cross-section, with all nine discs of the SCT End-cap having three nodes at the outer radius. Lines between these outer nodes form a rigid geodetic grid. Additional lines

extend to nodes at the inner rim of each disc, allowing the out-of-plane and in-plane deformations of each disc to be monitored. The number of inner nodes per disc is either 3 or 6, as detailed elsewhere [Gib03].

Assumptions of the model

A SIMULGEO model was written that described the grid geometry. The disc dimensions and relative separation of measurement nodes within the jewels were taken from engineering drawings [Wil03]. The simulated grid cross-section was equilateral, although the actual design is very slightly isosceles due to engineering constraints.

Each disc was defined in a distinct co-ordinate frame. The frame was defined by the 3 outer nodes of the disc and was free in all 6 external DoF. Relative movement of the outer nodes was allowed, by releasing the radial co-ordinates of the three outer nodes in the model. A small lateral translation of any outer node with respect to the reference frame was therefore interpreted in the model as being due to a lateral translation and/or axial rotation of the disc.

All grid nodes were defined as extended objects to model the FSI jewels. The orientation of the jewels within the local frame of the disc was completely constrained. The FSI components on the opposite sides of the disc at the same grid node location were rigidly fixed to each other and defined one grid node.

The jewels at the inner radius of type-A were completely constrained in the local R and RΦ co-ordinates but free in Z, within the disc. The local Z co-ordinate corresponds to an out-of-plane deformation of the disc. The jewels of type-B were completely constrained in R and free in RΦ and Z, in the frame of the disc.

One disc was defined as the reference, by constraining absolutely the 6 DoF to the global co-ordinate frame. All internal DoF of the reference disc were free.

A 1 μm FSI measurement precision was assumed on all GLIs. Less than 1 minute was required to read the model into SIMULGEO and the calculation took less than 1 second, on a 1.6 GHz Pentium® 4 processor and consumed 230 kbytes of memory.

5.4.7 Simulation results and discussion

The grid simulation yielded identical precision estimates on jewels of the same type within a disc as expected by symmetry. The calculated translational precisions on the X and Y co-ordinates of each disc were also identical, as were the rotational precisions about X or about Y.

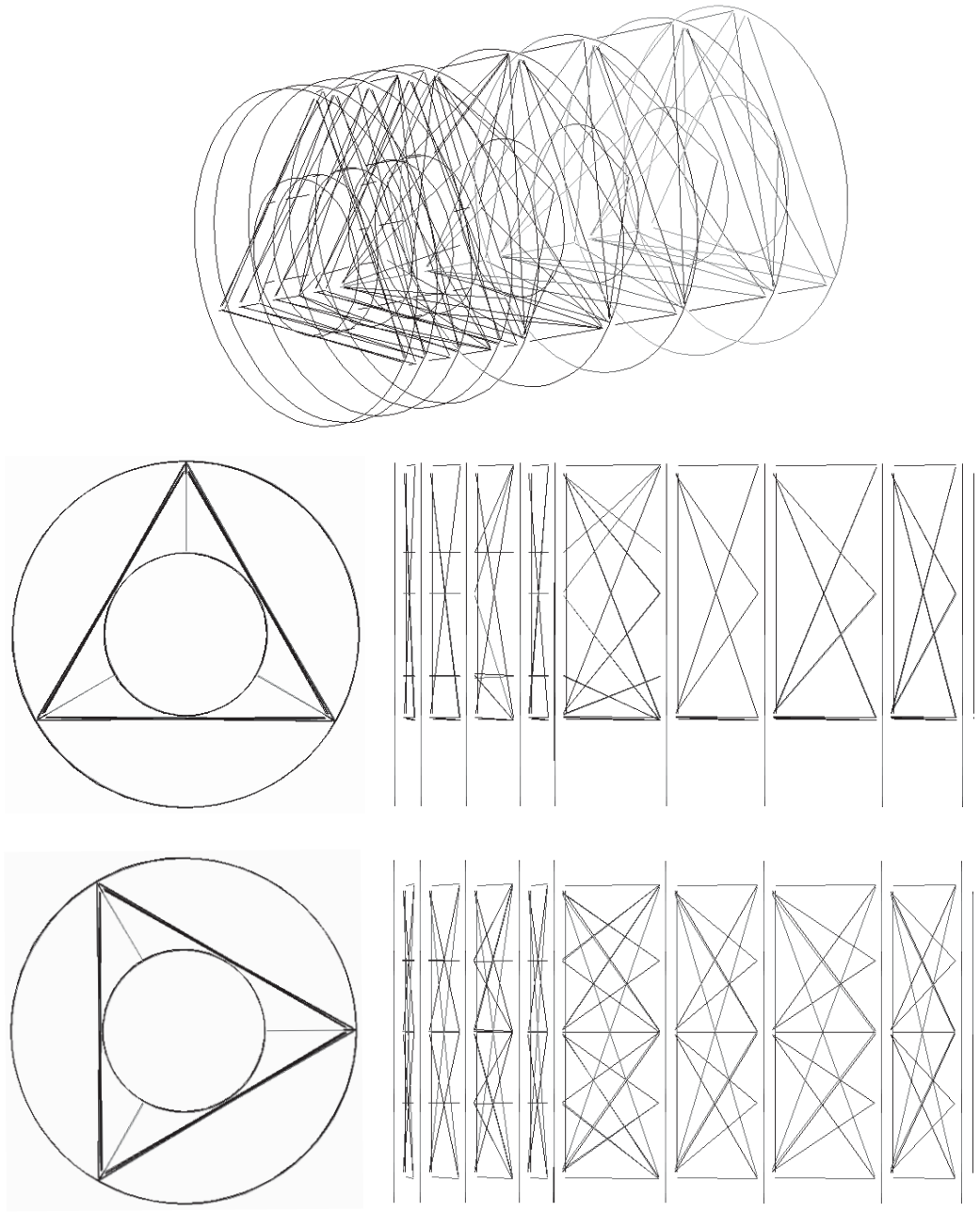

Figure 5.23: *The 165 length measurement lines of the FSI alignment grid for the nine discs of the SCT End-cap. A three dimensional rendering of the SCT End-cap is shown at the top, a side view is shown in the middle and a plan view is shown at the bottom. The cross-sectional views show that the triangular grid geometry fits neatly within the annular cross-section of the End-cap discs.*

The disc defined as the reference was systematically varied in repeated simulations and the calculated translational and rotational precisions are plotted in Figure 5.24.

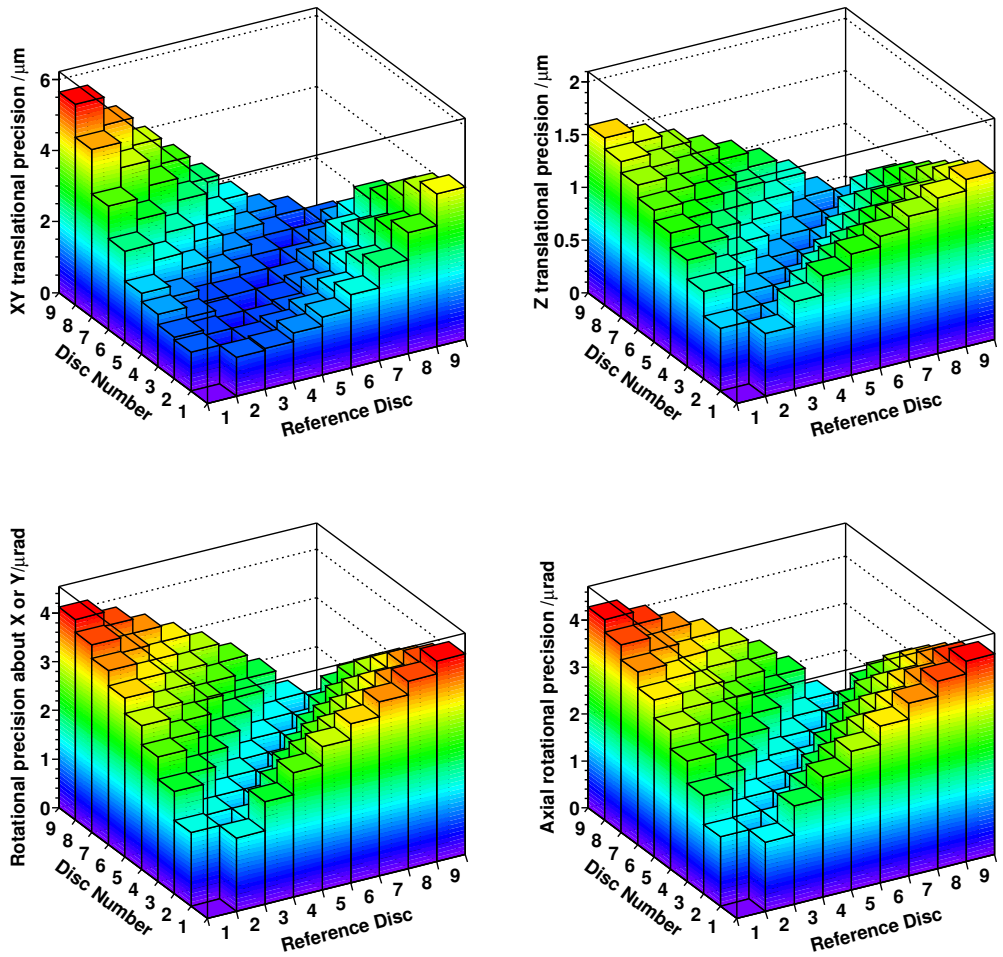

Figure 5.24: *Simulated precisions on the external degrees of freedom of the SCT End-cap discs. The precisions are plotted versus each disc, for nine simulations in which each disc was systematically defined as the reference.*

The geodetic grid performs extremely well, determining the transverse XY translation of disc 9 with respect to disc 1 to 5.6 μm. Notably, in the model where disc 9 is the reference, the XY translation of disc 1 with respect to disc 9 is 4.3 μm. This is due to the first few discs, nearest to the interaction point (1-5), being closer together in Z than the farthest discs (6-9). The angular

precision degrades more rapidly when the discs are closer in Z, as observed in the lower plots of Figure 5.24. When disc 1 is the reference, the translational error therefore increases more rapidly with distance from the reference disc, than when disc 9 is the reference. This highlights the benefit in choosing a reference frame in which the majority of nodes are the most stable. For example if the central disc is held reference, the XY translational precisions on discs 1 and 9 improve to 1.6 μm and 2.6 μm respectively.

The equivalent RΦ precision of a point on a disc with respect to the reference frame was estimated by summing in quadrature the XY translational and axial rotational precisions, accounting for the radial co-ordinate of the point, as plotted in Figure 5.25. The plot is for the worst case where disc 1 was the reference.

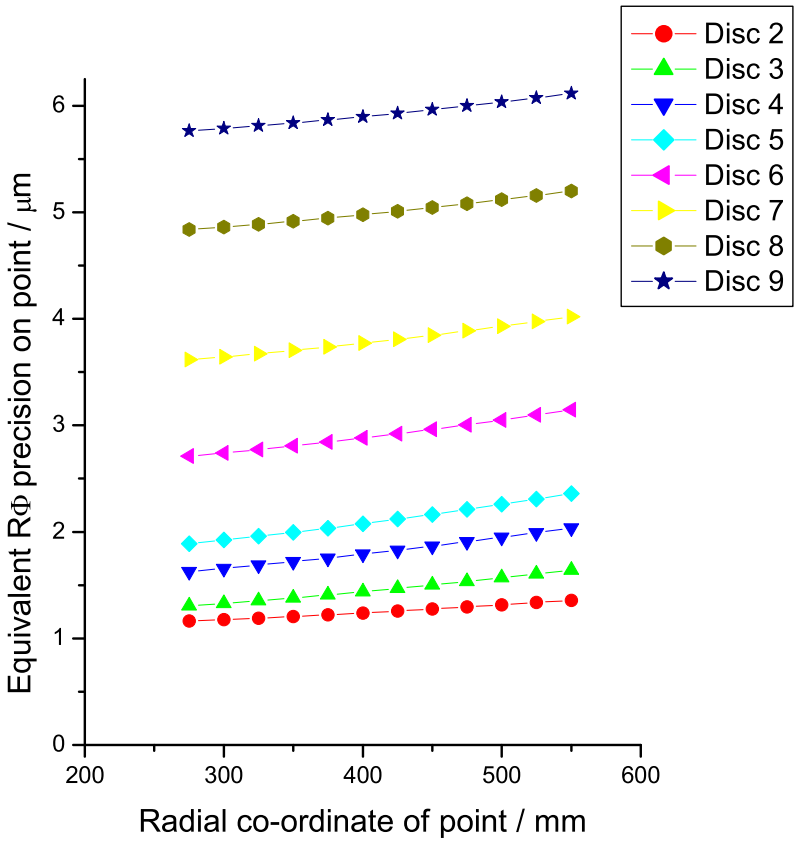

Figure 5.25: *Equivalent RΦ precision on a point at a particular radius within each disc with respect to disc 1, which defined the reference.*

Notably, the equivalent RΦ precisions are well within the 12 μm ATLAS requirement, especially since only four adjacent discs can be traversed by any

one track. The FSI system should perform very well, if the above assumptions prove to be true.

The precisions on the internal DoF of each disc are plotted in Figure 5.26. Due to the varying grid configuration not all discs have all node types, so only the measured DoF are plotted.

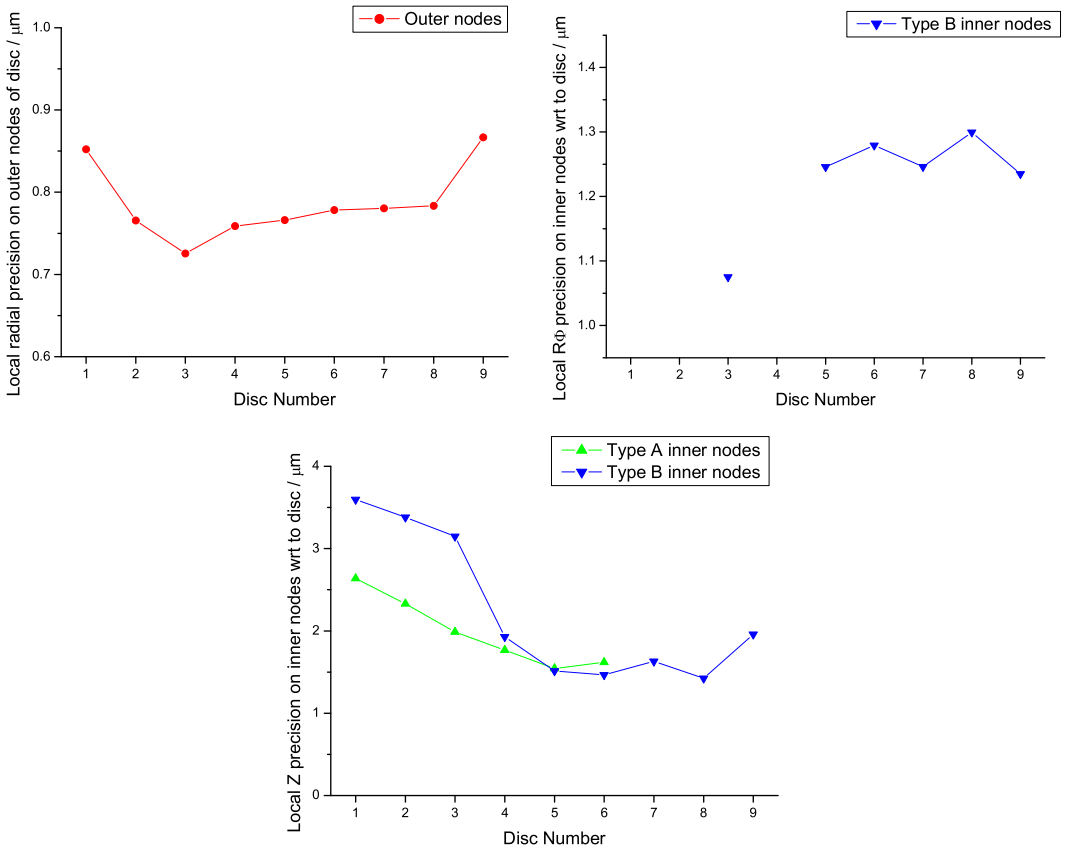

Figure 5.26: *Simulated precisions on the internal degrees of freedom of the SCT End-cap discs. The local precision on the grid nodes with respect to the disc frame are plotted versus disc number. All internal DoF precisions were independent of which disc defined the reference in the model.*

Notably, all precisions on the internal DoF are well within the ATLAS requirements and should help to determine the deformations of each disc with excellent precision. The small variation in precisions is due to the varying separation in Z between discs and number of GLIs incident on a particular node. For example, the precisions on the outer nodes of disc 1 and 9 are slightly worse, because the GLIs measure from only one side of those discs.

5.4.8 Conclusions

The SCT End-cap Grid simulations show the precision requirements of [ATL97] are met on the FSI nodes, with the above assumptions. These calculations indicate that the alignment system will be able to monitor displacements of the nodes on the last disc with respect to the first, to a precision of less than 7 μm for lateral displacements, and to less than 2 μm in Z.

5.5 ATLAS FSI System Implementation

The simulations in this chapter indicate that the FSI system should be capable of precisely reconstructing the displacements of nodes attached to the ATLAS SCT. The ultimate aim however, is to determine the SCT module co-ordinates, through the combination of data from the FSI system and from tracks, and through knowledge of the SCT support structure. This section outlines how this task might be broken down and developed into a coherent alignment strategy. The aim is to provide a suggested but not exhaustive guide for future work.

FSI length measurements The multiple, simultaneous length measurements made by the ATLAS FSI system, will be the first step in determining the shape of the SCT. The on-detector FSI hardware is currently being installed. Developments continue for the off-detector FSI laser system and read-out hardware. The prototype FSI system software should be developed to allow signals from the 842 GLIs in ATLAS to be reliably processed to generate precise length measurements.

Grid node reconstruction These precise length measurements will be combined to reconstruct the ATLAS FSI grid node co-ordinates. The SIMULGEO models described in this chapter could be built upon and additional software could be developed to address some limitations of SIMULGEO in defining complex mechanical constraints. The following features would aid the reconstruction of ATLAS FSI grids:

- Identification and removal of outlying/biased FSI measurements.
- Ability to define the average position between a set of grid nodes.
- Ability to define sophisticated correlations of node positions (for example, with eigenmode deformations of the support structure, as discussed below).

The computational optimization of SIMULGEO [Bru98] indicates each section of the ATLAS grid should be reconstructed on short time-scales.

If further constraints are added to the models, these time-scales would increase with the cube of the number of parameters [Bru98]. If necessary, at the cost of some precision, the grid could be broken down into the natural sub-grids and solved independently, for later recombination in a simple master model.

Node interpolation by shape fitting The reconstruction software will produce grid node positions with associated errors, which must be interpolated to find the module co-ordinates. Shape fitting algorithms are being developed to fit models of the detector support structure to the grid node positions.

The current grid simulation software only permits simple mechanical constraints. The shape fitting algorithms should allow the expected deformations of the support structures, to improve the precision and reliability of the method and exclude unphysical solutions. FEA models and analytical models that parameterize the detector shape should be investigated.

Calibration A method of relating the fitted support structure shapes to the module co-ordinates should be developed. These relations, which may be considered as module offsets from the fitted shape, must be determined. An X-ray survey [Dou01] of the SCT should determine the initial module co-ordinates and a simultaneous FSI scan should enable calibration of these module offsets. The precision with which the alignment system can predict the module co-ordinates should improve with time, as track data is used to refine the calibration.

Additionally, the node offsets within the FSI jewels should be calibrated by the methods in Chapter 4.

Validation Each step outlined above should be tested both in software and via an experimental programme based on FSI measurements and simultaneous independent measurements (e.g. ESPI) of a prototype SCT support structure. The tests would allow the shape reconstruction models to be tuned to match real measurements. The effects of temperature variations on module positions should also be investigated.

Integration with tracks The spatial frequency information and time dependencies of particle tracking data and FSI system data complement each other very well. Track data accesses the high spacial frequency eigenmode deformations, over long time-scales, and FSI data constrains the low spatial frequency modes, on short time-scales. The best alignment precision should be achieved by combining these data.

In particular, the alignment system will be critical to understanding and correcting for sagitta deformations and short term distortions, that otherwise degrade purely track based alignment methods. Combined with tracking data, the FSI system should also aid alignment of the Pixels and TRT, which do not have separate alignment systems. The SCT alignment system should therefore be a valuable tool in aligning the ATLAS Inner Detector.

5.6 Conclusion

This chapter has presented detailed simulations of the FSI geodetic grids that will monitor the ATLAS SCT. The simulations systematically evaluated the various grid sections. In general, the simulations indicate that the FSI system should be capable of reconstructing displacements of the grid nodes to well within the required precision.

Simulations of the SCT barrel grid indicate that the required precision should be achieved, if certain assumptions on the mechanical rigidity of the SCT support structure prove to be true. The rigidity of the end-flange was found to be an important constraint, which strengthened the grid rigidity and redundancy. Additionally, the local rotations of the FSI jewels must be less than about 100 μrad to have negligible effect on the reconstruction, which is similar to the maximum rotation estimated from FEA calculations. The strengths and weakness of the grid were investigated, by evaluating how the sequential addition of grid sections improves the precision. The importance of maintaining the closed loop of GLI measurements around one barrel was demonstrated.

An FSI grid for the SCT End-cap has been developed through detailed simulation studies. The grid design was optimized for rigidity and redundancy, while accounting for several conflicting constraints. Instructive grid design failures highlighted the strengths and weakness of the certain configurations and demonstrated the relevance of the global reference frame definition. Simulations of the optimized design indicate the grid nodes should be reconstructed to well within the required precision, provided certain mechanical constraints prove to be true.

The future challenge will be to combine data from the FSI system and from tracks to align the ATLAS Inner Detector. A programme to guide these future developments has been suggested.

Chapter 6

Comparative Study of Misalignment at CDF and ATLAS

6.1 Introduction

Tagging the flavour of b-jets is crucial to the identification of many Higgs and SUSY channels expected at the LHC. Misalignment of the ATLAS inner detector degrades the tracking resolution and hence the lifetime-based b-tagging performance. During ATLAS commissioning the main challenge faced in aligning the detector will be to understand systematics caused by sagitta deformations and detector instabilities. The ATLAS SCT alignment system should quicken the commissioning phase by helping to elucidate complex and time-varying detector deformations. The extent of these deformations is difficult to accurately predict from engineering constraints. In this chapter, misalignments observed at a current experiment during its commissioning phase are examined and the implications for ATLAS are investigated. The aim is to provide an indicative estimate of the misalignments that are to be tackled at ATLAS and to emphasize the need for the SCT alignment system.

6.2 CDF II and Alignment

Due to similarities with ATLAS and as the current best hadron collider experiment at the high energy frontier, the Collider Detector at Fermilab (CDF) was chosen as the current experiment, with which to study the effects of misalignment. The upgraded detector, called CDF II [CDF96b], resumed data taking in autumn 2001 to study $\mathrm{p\overline{p}}$ collisions at $\sqrt{s} = 1.96$ TeV provided by

the Tevatron. This current data taking period, called Run IIa, is the source of the CDF II data in this chapter [Pau03].

The CDF II tracking system relevant to this chapter is briefly summarized.

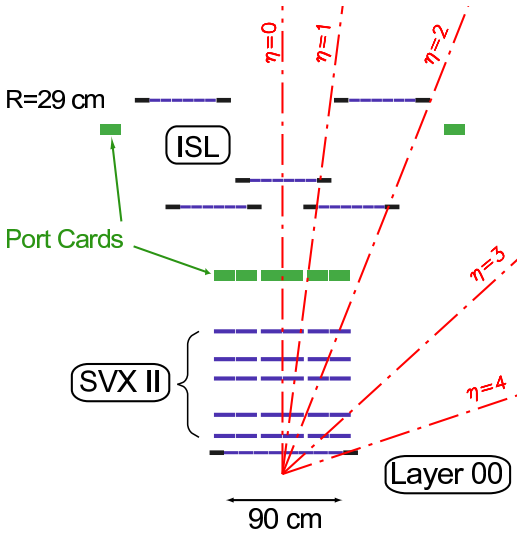

Figure 6.1: *rz cross-section of half of the CDF II silicon tracking system. The z scale is compressed.*

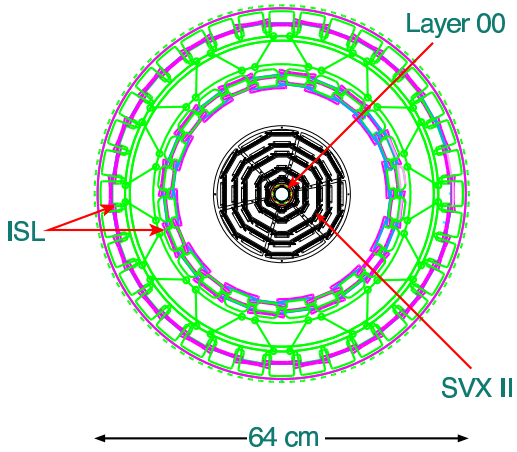

Figure 6.2: *View in the rφ plane of the CDF II silicon tracking system.*

The CDF II silicon tracking system (SVX) [Sil00] is shown in Figures 6.1 and 6.2. This system provides up to 8 precision hits per track from the three subsystems: a single layer of single-sided silicon detectors (Layer 00) is mounted directly onto the beam-pipe to improve the impact parameter resolution and b-tagging performance; the Silicon Vertex Detector (SVX II) comprises 5 concentric layers of double-sided silicon ladders; and the Intermediate Silicon Layer (ISL) comprises 3 concentric but not hermetic layers of double-sided silicon ladders. A drift chamber called the Central Outer Tracker (COT) completes the CDF II inner tracking system [Bur01].

The relative alignment of the tracking subsystems is monitored by a RASNIK system [Sal98, Gol02]. The global alignment can be actively controlled with micro motors, called *inchworms*, based on piezoelectric actuators. Time periods in which the detector was particularly stable were identified with the RASNIK alignment system; the observed relative detector movements were < 5 μm in a day. On more typical days, movements of ~ 30 μm were observed, corresponding to changes in the solenoid current. The tracking system returned to its nominal position during data taking, although some creep of the nominal position was observed over the course of a month.

The CDF II alignment system plays an important role in identifying stable detector periods and flagging when new global detector alignment constants need to be defined, as provided by track analysis [McN01]. In this study, data from CDF II were reconstructed with two different sets of alignment constants corresponding to those from an initial mechanical survey of the tracking system and from the best alignment constants as of September 2003. These data were compared to check the impact of initial misalignments on track parameter resolutions.

6.3 Momentum Resolution Degradation

6.3.1 Introduction

In general, misalignments degrade all of the 5 measured helical track parameters: d_0, z_0, ϕ_0, $\cot\theta$, Q/p_T. In this section, the degradation due to misalignment on the transverse momentum, p_T, is studied. The $1/p_T$ resolution can be parameterized as:

$$\sigma\left(\frac{1}{p_T}\right) = \sqrt{A^2 + (B/p_T)^2} = A \oplus B/p_T \tag{6.1}$$

where $p_T = p\sin\theta$ (GeV)[1]. The two terms A and B correspond respectively to the contribution from the intrinsic detector resolution and the contribution from multiple scattering. These terms are different for CDF II and ATLAS, due to the different detector designs. The ATLAS silicon trackers (pixels & SCT) provide better intrinsic detector resolution than the CDF II silicon system but contain more material[2], resulting in a larger multiple scattering term for ATLAS than for CDF II. These differences must be accounted for when studying the effects of misalignment in the different experiments.

The aim of this study was to determine the difference in the $\sigma(1/p_T)$ parameterization at CDF II between the initial and best alignments, in order to map the effects of misalignment at CDF II onto ATLAS. This was achieved by comparing the mass resolution of a narrow resonance in CDF II as reconstructed with either the initial survey or best alignment constants. Any difference in the measured width of the invariant mass peak was assumed to be the result of the initial misalignment.

[1] In this chapter mass, momentum and energy are quoted in natural units: $c = \hbar = 1$.

[2] The radiation length for CDF II silicon system is roughly $0.1\ X_0/\sin\theta$ for $|\eta| < 1$ and about twice this for $1 < |\eta| < 2$ [Sil00]. The radiation length expected for the ATLAS silicon trackers (pixels + SCT) is around $0.2\ X_0$ for $|\eta| < 1$ rising to around $0.6\ X_0$ at $|\eta| = 2$ [Haw04]

The $1/p_T$ resolution for the SVX-CTC system of CDF was parameterized as [CDF94]:

$$\sigma\left(\frac{1}{p_T}\right) = \sqrt{0.0009^2 + (0.0066/p_T)^2} \tag{6.2}$$

where $p_T = p\sin\theta$ (GeV). CDF II was designed to have $\delta p_T/{p_T}^2 < 0.1\%$ GeV [Bur01].

The measured curvature for stiff tracks should be degraded the most by misalignments, whereas multiple scattering dominates at lower track momentum. It was necessary to choose a resonance for which the tracks left by the decay products were of sufficient momenta for the misalignment induced degradation to be observable.

The resonant channel $J/\psi \to \mu^+\mu^-$ was considered, but does not meet this requirement because the low mass of the J/ψ, $m_{J/\psi} = 3096.87 \pm 0.04$ MeV [PDG02], leads to decay products with low momenta. Hence, for a J/ψ produced at rest the two terms on the RHS of Equation 6.2 differ: $0.0009^2 \ll (0.0066/\frac{m_{J/\psi}}{2})^2$, so the J/ψ mass resolution at CDF II is dominated by multiple scattering.

The resonant channel $\Upsilon(1S) \to \mu^+\mu^-$ was studied instead, because $m_{\Upsilon(1S)} = 9460.30 \pm 0.26$ MeV [PDG02] so for a decay at rest, both terms are of the same magnitude: $0.0009^2 \approx (0.0066/\frac{m_{\Upsilon(1S)}}{2})^2$. The $1/p_T$ degradation due to misalignment should therefore be observable in the $\Upsilon(1S) \to \mu^+\mu^-$ channel at CDF II.

6.3.2 Υ(1S) Resonance at CDF II

A data sample from CDF II that contained $\Upsilon \to \mu^+\mu^-$ events was kindly provided for this analysis [Pau03]. In this data sample, the tracks were reconstructed with the best available alignment constants for the event selection procedure. These tracks were then refitted with either the alignment constants from the initial mechanical survey or from the current[3] best alignment constants. These two data samples were then sorted by run and event numbers and duplicate tracks from the combinatorial selection process were removed. Only tracks that were present in both data samples were included in the analysis. Four or more $r\phi$ silicon hits were required for each track. A p_T cut of 3 GeV was present in the data sample as shown by the p_T distributions in Figure 6.3. No further p_T cuts were made.

The invariant mass distribution of pairs of oppositely charged muons from each event is plotted in Figure 6.4, in which the first three Υ S-wave resonances are apparent. A Gaussian + linear background fit to the $\Upsilon(1S)$ peak, over the

[3]The data were processed in September 2003.

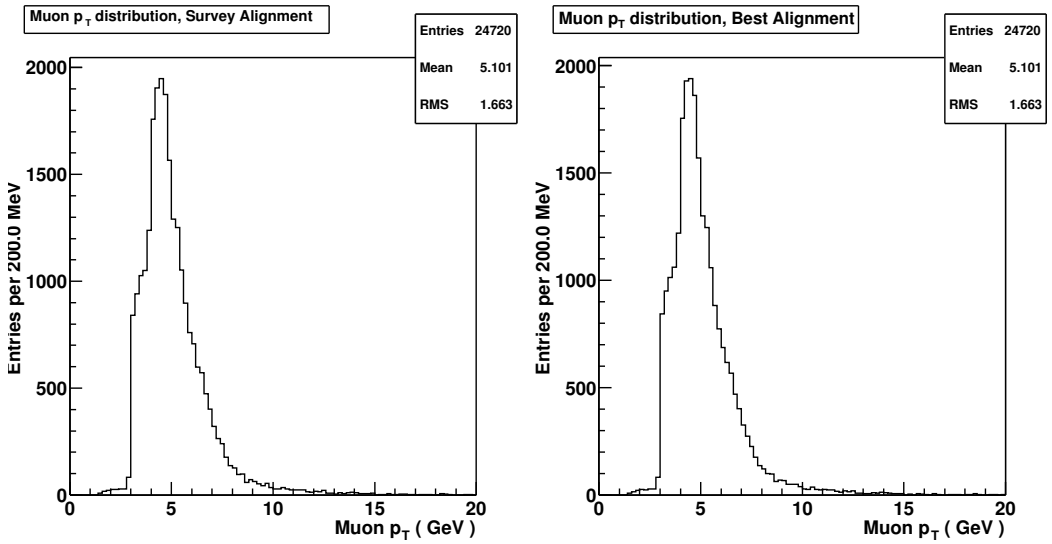

Figure 6.3: *Muon transverse momentum distributions for CDF run IIa data for initial mechanical survey and best alignments. The plots only include pairs of oppositely charged muons that left at least 4 rϕ silicon hits and have an invariant mass between 9.1 GeV and 9.8 GeV.*

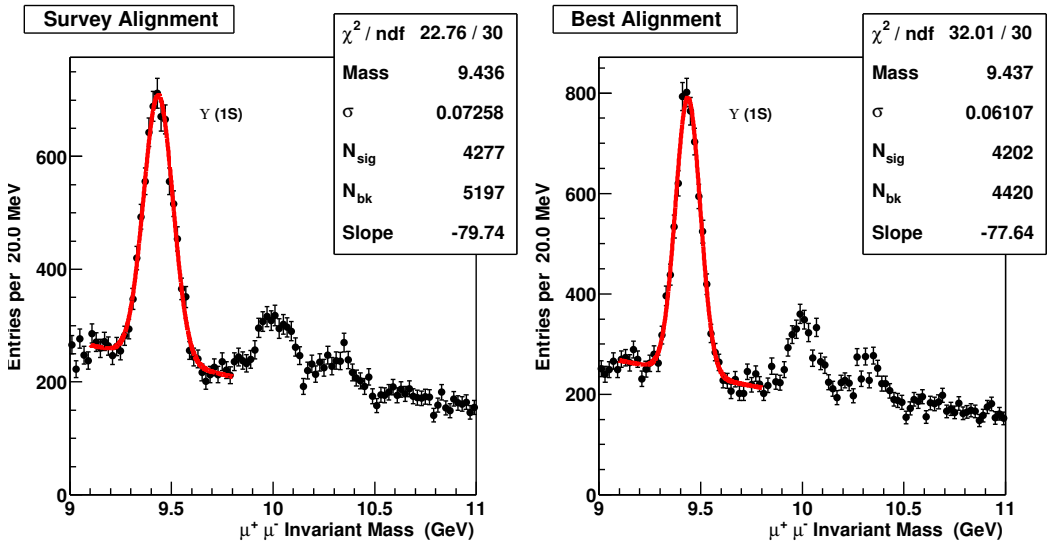

Figure 6.4: *Υ(1S) invariant mass peak reconstructed from CDF run IIa data with initial survey and best alignments. The Υ(2S) and Υ(3S) are also visible. The change in the mass resolution, σ, is indicative of the misalignment.*

dimuon invariant mass range of 9.1-9.8 GeV yielded a width of 72.6 MeV for the initial alignment and 61.1 MeV for the best alignment.[4] The Υ(1S) mass in either case is lower than the world average mass, primarily because no vertex constraint was applied and also no corrections for the momentum scale and internal bremsstrahlung were applied as had been in other work [CDF96a].

6.3.3 Monte Carlo Tuning Method

The difference observed in the Υ(1S) widths indicates the difference in misalignment between the initial mechanical survey and the best alignment constants. In order to apply the observed effect to ATLAS, the ratio of resolution terms in the $\sigma(1/p_T)$ parameterization of Equation 6.1 was determined by tuning a Monte Carlo simulation of the Υ(1S) dimuon decay to match the observed peak widths.

The measured track curvature, $C \approx \frac{0.3B}{2p_T} \propto \frac{1}{p_T}$, should be degraded by approximately Gaussian errors, so $\frac{1}{p_T}$ was smeared in the Monte Carlo. Misalignments were assumed to contribute only to the detector resolution term, A in Equation 6.1. The multiple scattering terms were assumed identical $B_S = B_B = B$, where the subscripts S and B denote the survey and best alignments. Defining the difference in detector resolution terms to be due to misalignment, $A_{\mathrm{AL}} = A_S - A_B$, the inverse p_T resolutions for either survey or best alignment are:

$$\sigma_S\left(\frac{1}{p_T}\right) = A_S \oplus B_S/p_T = (A_B + A_{AL}) \oplus B/p_T \tag{6.3}$$

$$\sigma_B\left(\frac{1}{p_T}\right) = A_B \oplus B_B/p_T = A_B \oplus B/p_T \tag{6.4}$$

The ratio $F_{S/B}^{P_T} = A_S/A_B$ was sought.

Simulation Chain

Signal generation The Monte Carlo event generator, PYTHIA 6.203 [Sjo01], was configured for a $p\overline{p}$ collider at $\sqrt{s} = 1.96$ TeV to replicate the Tevatron. 10,000 events were simulated in which a Υ(1S) produced in gluon-gluon fusion was forced to decay via $\mu^+\mu^-$.

ATLFAST These events were passed through the ATLAS fast simulation, ATLFAST [Ric98], with the muon track smearing turned *off*. ATLFAST

[4] Both widths reduced very slightly if a vertex constraint was also applied, but with negligible effect on the following analysis because only the width ratio matters to first order.

formed part of the analysis chain only to apply muon level 1 trigger cuts of $p_T > 1.4$ GeV and $\eta < 2.5$, and to allow access to the unsmeared track data in the ATLAS release 6.6.0 Athena framework, in which the data were analysed.

Custom track smearing The track smearing was performed by a simple custom algorithm that smeared the inverse transverse momentum according to the parameterization in Equation 6.1. A further cut of $p_T > 3.0$ GeV was performed to simulate the cut present in the CDF II p_T distributions in Figure 6.3.

Figure 6.5 shows the result of smearing the inverse track momenta according to the CDF parameterization given in Equation 6.2. A parametrized fit to the difference in $1/p_T$ between the smeared and unsmeared track demonstrates the smearing proceeded as expected. The invariant mass distribution of pairs of oppositely charged muons in the lower two plots show the Υ(1S) peak reconstructed from unsmeared and smeared tracks. A Gaussian fit to the lower right plot yields a Υ(1S) mass resolution of 52.6 MeV for the CDF parameterization constants (A=0.0009, B=0.0066). This does not match the mass resolution of 61.1 MeV found for the best alignment at CDF II in Section 6.3.2, because the parameterization for the CDF II detector should be slightly different than for CDF, due to more material, and because no vertex constraint was applied.

The Monte Carlo was tuned by repeating the track smearing for different values of A while retaining a constant value of B. The fitted Υ(1S) mass resolution in each case is plotted in Figure 6.6 as a function of the smearing parameter A. These results were interpolated by a second order polynomial fit, to evaluate the two values of A for which the simulated Υ(1S) mass resolution should match that found in CDF II. Finally the smearing was repeated for each of these two values, given in Table 6.1, to produce the simulated Υ(1S) peaks in Figure 6.7, for which the mass resolutions are compatible with those from CDF II data in Figure 6.4. The misalignment induced degradation factor is estimated from Table 6.1 as $F_{S/B}^{P_T} = A_S/A_B \approx 1.4$. The implications for ATLAS are discussed in Section 6.4.

Alignment	CDF II Υ(1S) mass resolution (MeV)	Simulation matched for smearing parameters:		Misalignment Factor A_S/A_B
		A	B	
Survey	72.6	0.00183	0.0066	1.4
Best	61.1	0.00134	0.0066	

Table 6.1: *Smearing parameters for which the simulated Υ(1S) mass resolutions matched those from CDF II data for the survey and best alignments.*

6.3.4 $1/p_T$ Comparative Method

A method was devised to estimate the difference, $A_{AL} = A_S - A_B$, directly from the CDF II data, to corroborate the Monte Carlo tuning method. The method

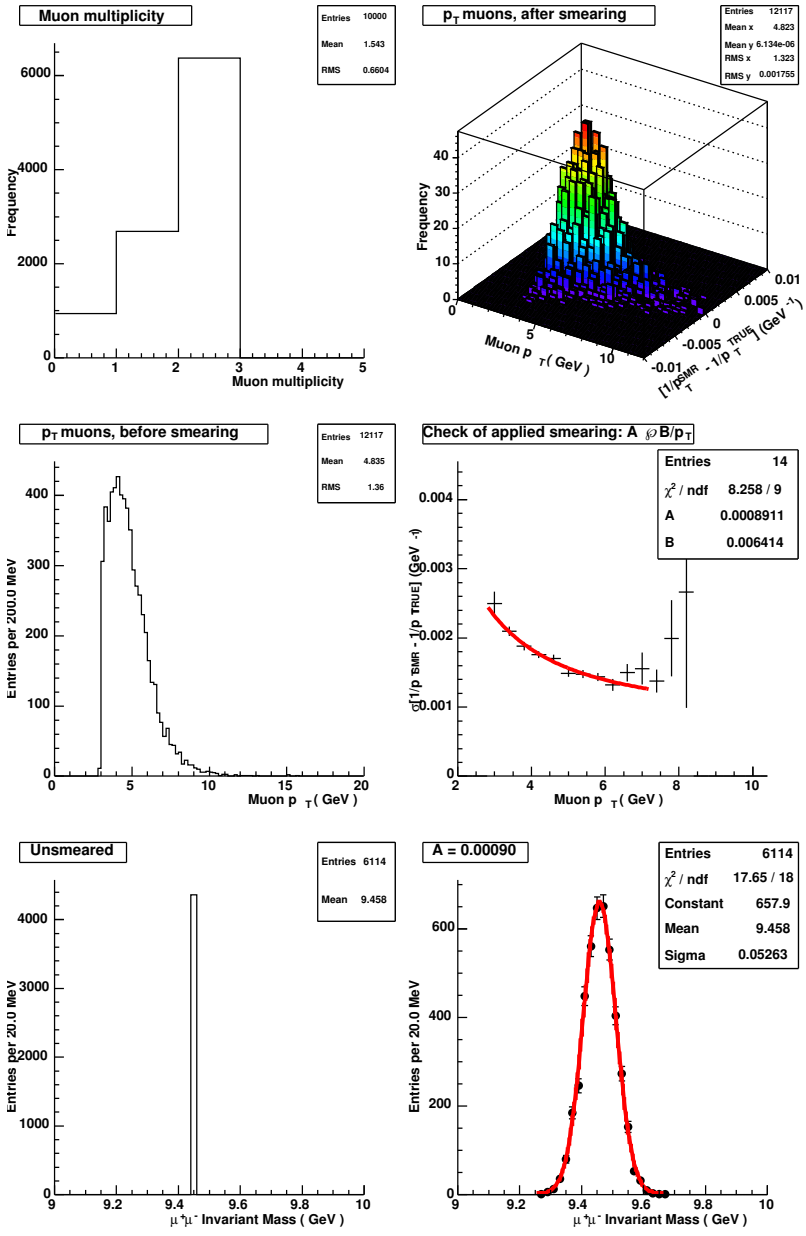

Figure 6.5: *Reconstruction of* $\Upsilon(1S) \rightarrow \mu^+\mu^-$ *as simulated with and without track* $1/p_T$ *smearing, according to the CDF* $1/p_T$ *resolution parameterization (A=0.0009, B=0.0066).*

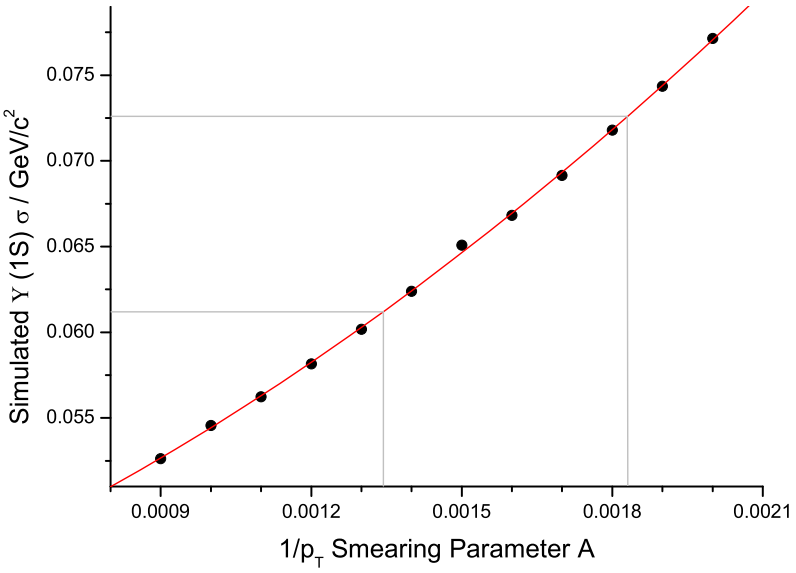

Figure 6.6: *The simulated fitted mass resolution of the Υ(1S) as a function of the smearing parameter A. The parameter B was 0.0066 in all simulations. The gray lines highlight the two values of A, given in Table 6.1, for which the simulation matched CDF II data for survey and best alignments.*

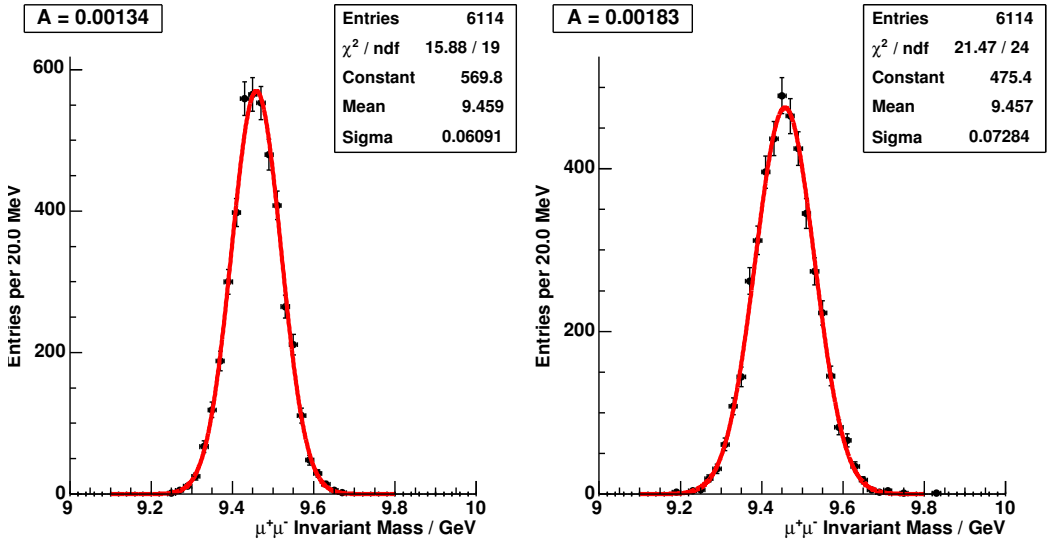

Figure 6.7: *Reconstructed* $\Upsilon(1S) \rightarrow \mu^+\mu^-$ *for track smearing parameters estimated from Figure 6.6. The fitted mass resolutions closely match those observed in CDF II data for the survey and best alignment data samples.*

by which the CDF II data samples were generated enabled a direct comparison of the two tracks that corresponded to the same particle, but that were reconstructed with the two different sets of alignment constants. The effective $1/p_T$ smearing for the tracks in the alignment data samples was assumed to have proceeded according to Equations 6.3 and 6.4, such that:

$$\frac{1}{p_T^{\mathrm{SURVEY}}} = \frac{1}{p_T^{\mathrm{TRUE}}} + (A_B + A_{AL})G_1(0,1) + \frac{B}{p_T^{\mathrm{TRUE}}}G_2(0,1) \quad (6.5)$$

$$\frac{1}{p_T^{\mathrm{BEST}}} = \frac{1}{p_T^{\mathrm{TRUE}}} + A_B G_1(0,1) + \frac{B}{p_T^{\mathrm{TRUE}}}G_2(0,1) \quad (6.6)$$

where $G_n(0,1)$ is a Gaussian distributed random number of $\mu = 0$, $\sigma = 1$. On taking the difference, (6.5) - (6.6), in the inverse p_T of the two tracks for a given particle, all terms should cancel[5] except for the residual misalignment term, A_{AL}:

$$\frac{1}{p_T^{\mathrm{SURVEY}}} - \frac{1}{p_T^{\mathrm{BEST}}} = A_{AL}G_1(0,1) \quad (6.7)$$

$$\sigma\left(\frac{1}{p_T^{\mathrm{SURVEY}}} - \frac{1}{p_T^{\mathrm{BEST}}}\right) = A_{AL} \quad (6.8)$$

This difference in the inverse p_T of particle tracks reconstructed with each set of alignment constants is plotted in Figure 6.8, for all tracks that contributed[6] to the fitted Υ(1S) peaks in Section 6.3.2. The width of a Gaussian fit to this $\frac{1}{p_T^{\mathrm{SURVEY}}} - \frac{1}{p_T^{\mathrm{BEST}}}$ distribution provides, from Equation 6.8, an estimate of $A_{AL} \approx 0.00048$. This estimate closely matches the equivalent estimate from the Monte Carlo tuning method of $A_{AL} = A_S - A_B \approx 0.00049$, from Table 6.1.

The assumptions in the above parameterization imply that there should be no p_T dependence in $\sigma(\frac{1}{p_{\mathrm{T}}^{\mathrm{SURVEY}}} - \frac{1}{p_{\mathrm{T}}^{\mathrm{BEST}}})$, but a significant p_T dependence was observed, as shown in Figure 6.9. This p_T dependence is thought to arise either from correlations between the misalignment and the detector resolution and multiple scattering terms, or from a systematic (rather than random) misalignment such as the sagitta deformation suggested below. The p_T dependence does not invalidate the Monte Carlo Υ(1S) method or the $1/p_T$ comparative method, which both integrate over all tracks to average out the p_T dependence and to provide a general and indicative measure of the misalignment.

[5] Correlations are assumed between the different alignment data samples for the multiple scattering and detector resolution terms. A simple correlated Monte Carlo study confirmed only the alignment term remains.

[6] Only oppositely charged muon pairs having an invariant mass of 9.1-9.8 GeV were included in the plots of Figures 6.8, 6.9 and 6.10.

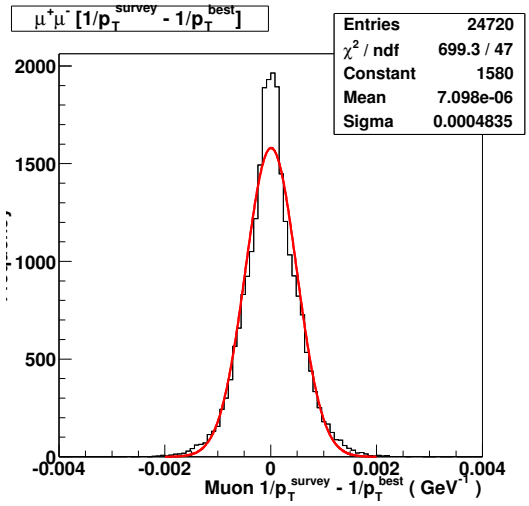

Figure 6.8: *CDF run IIa data. Difference between the inverse transverse momentum of each track, when reconstructed with the mechanical survey and best alignment constants. The width of this distribution σ=0.00048 GeV^{-1} is indicative of the misalignment.*

An interesting misalignment effect is revealed by independently plotting the $1/p_T$ difference for μ^+ or μ^- tracks, as in Figure 6.10. The distributions for μ^+ or μ^- are individually asymmetric, but the combined distribution is symmetric. This observation suggests that some of the misalignment is due to a sagitta deformation, perhaps caused by the incremental relative rotation of tracking layers about the beam axis. A deformation such as this would create a curvature error of opposite sign for the oppositely charged tracks, producing the observed asymmetries in the measured track momentum. Similar sagitta deformations in ATLAS should be identified and constrained by the SCT alignment system.

6.3.5 Conclusion

The effect of misalignments at CDF II on the transverse momentum resolution has been studied. On examining the $\Upsilon \rightarrow \mu^+\mu^-$ channel and by tuning the track smearing parameterization in Monte Carlo to match the CDF II data, a degradation factor of 1.4 was found between the initial and best alignments. The Monte Carlo study was corroborated directly from CDF II data using a $1/p_T$ comparative method. This second method also indicated sagitta deformations were a substantial source of the misalignment at CDF II during the commissioning phase.

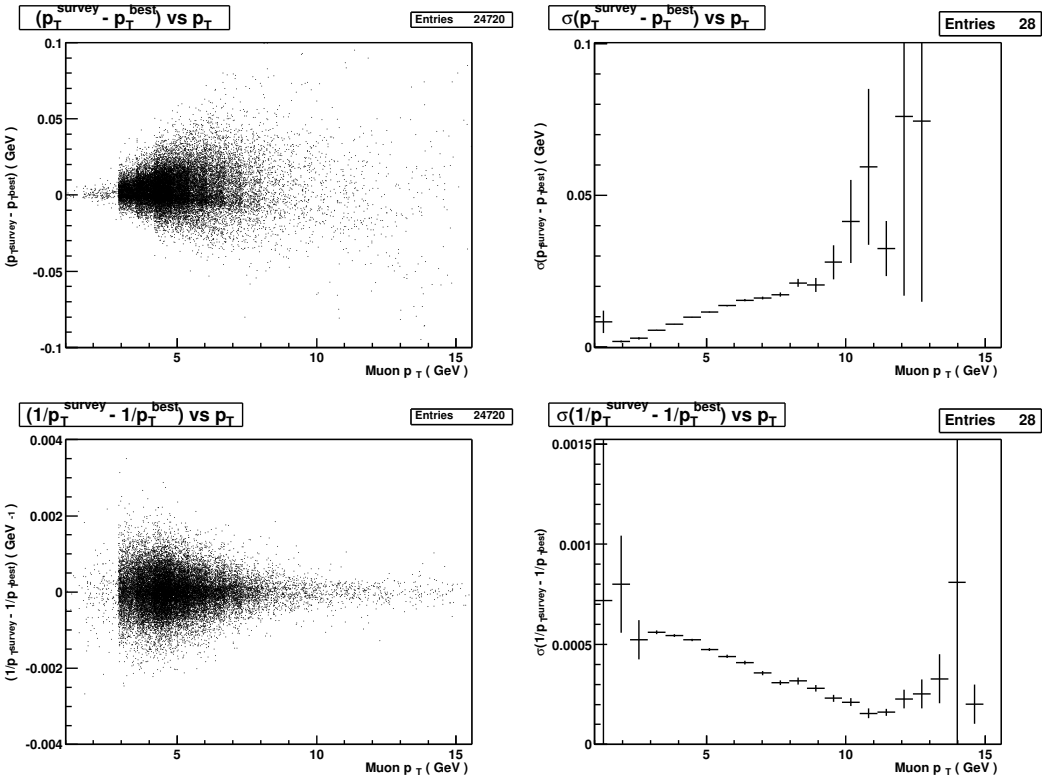

Figure 6.9: *CDF run IIa data. The difference in p_T and in $1/p_T$ between survey and best alignments, as a function of p_T. A set of Gaussian fits in slices of p_T reveals that $\sigma(\frac{1}{p_{\mathrm{T}}^{\mathrm{SURVEY}}} - \frac{1}{p_{\mathrm{T}}^{\mathrm{BEST}}})$ is p_T dependent, rather than p_T independent as expected from the simple parameterization assumed.*

6.4 Effect of Momentum Resolution Degradation

Effect of misalignment on $1/p_T$ resolution and on ATLAS physics

The design of the ATLAS inner detector has changed from that presented in the Physics TDR, resulting in some loss of detector performance [ATL97, Haw04]. Further performance losses due to detector misalignments could have a considerable impact on ATLAS physics, particularly during detector commissioning.

The $1/p_T$ resolutions given below correspond to the ATLAS inner detector geometry as given in the Physics TDR and to the updated geometry in recent

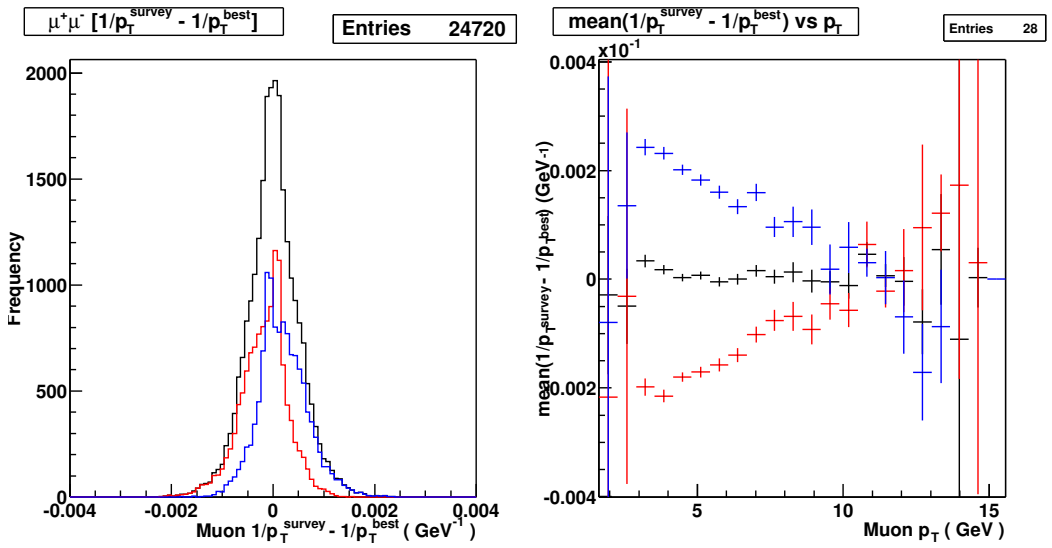

Figure 6.10: *CDF run IIa data. The difference in $1/p_T$ between the survey and best alignments, for: μ^+ in blue, μ^- in red, and all muons in black. The opposite asymmetries in the individual μ^+ and μ^- distribution indicate a sagitta misalignment.*

Monte Carlo simulations (Data Challenge 1) [Gar03a] :

$$\sigma^{\mathrm{LIN}}_{\mathrm{TDR}}\left(1/p_T\right) = 0.0004 + 0.0147/p_T \quad \text{TDR geometry} \tag{6.9}$$

$$\sigma^{\mathrm{LIN}}_{\mathrm{DC1}}\left(1/p_T\right) = 0.0006 + 0.0180/p_T \quad \text{DC1 geometry} \tag{6.10}$$

for p_T in GeV.

Assuming that the ATLAS inner detector suffers a misalignment induced degradation factor of 1.4 in the resolution term, as observed during commissioning at CDF II, then the $1/p_T$ resolution during commissioning at ATLAS would be degraded to[7]:

$$\sigma_{\mathrm{COMS}}\left(1/p_T\right) \approx 0.0008(4) \oplus 0.0180/p_T \tag{6.11}$$

The multiple scattering term should dominate any misalignment effect in the $1/p_T$ resolution of ATLAS for $p_T \lesssim (0.0180/0.0006) = 30$ GeV. The effect of misalignments on the $1/p_T$ resolution in ATLAS should therefore be

[7]Note that the parametrization in Equation 6.10 is a linear sum of terms a and b/p_T, rather than the quadrature sum, as used in the ATLAS TDR and for the above CDF II analysis. In the asymptotic limits at low or high p_T, the parameterizations are approximately equivalent: $a + b/p_T = a \oplus \sqrt{2ab/p_T} \oplus b/p_T \approx a \oplus b/p_T$. The quadrature sum parameterization was taken in the following analysis, in which only the ratio of the a terms is relevant.

negligible for low p_T B-physics. The above study indicates that misalignments would severely degrade the resolution in higher p_T physics channels, but the extrapolation to high p_T is not necessarily valid, considering that the studied CDF II data only contained tracks of low p_T ($\lesssim 10$ GeV). However, this large degradation factor and associated loss in detector performance emphasizes the importance of reducing the misalignments in ATLAS as quickly as possible, with the aid of the SCT alignment system.

Detector misalignments are expected to have a significant effect on the impact parameter resolution, as discussed in Section 6.5, with resulting degradations in b-hadron lifetime measurements, the b-tagging performance and the efficiency of the B-physics trigger [Bai01].

6.5 Impact Parameter Degradation

6.5.1 Introduction

The long lifetime and high mass of b-hadrons causes tracks left by their decay products to have large impact parameters[8]. The displaced vertex may be exploited to tag the b-jet provided that the tracking detector can resolve the impact parameter with sufficient precision. The CDF II data sample which contained tracks reconstructed with different alignments, as described in Section 6.3.2, allowed investigations into the effect of misalignment on the impact parameter resolution, $\sigma(d_0)$, which is parameterized by:

$$\sigma(d_0) = \sqrt{a^2 + (b/p_T)^2} = a \oplus b/p_T \tag{6.12}$$

for $\sigma(d_0)$ in μm and p_T in GeV.

The two terms represent, a, the asymptotic resolution at infinite momentum and b, the dependence on multiple scattering at low momentum. The aims of this study were to find the ratio in the resolution term, a, due to misalignment and to apply this effect to ATLAS.

6.5.2 Impact Parameter Distribution

The impact parameter distribution was investigated for each CDF II alignment data sample. Tracks were selected by the method in Section 6.3.2, so that 4 or more $r\phi$ silicon hits were required for oppositely charged pairs of muon tracks that had an invariant mass of 9.1-9.8 GeV. The track impact parameters for

[8]The impact parameter of a track is the smallest separation between the fitted track helix and the primary vertex.

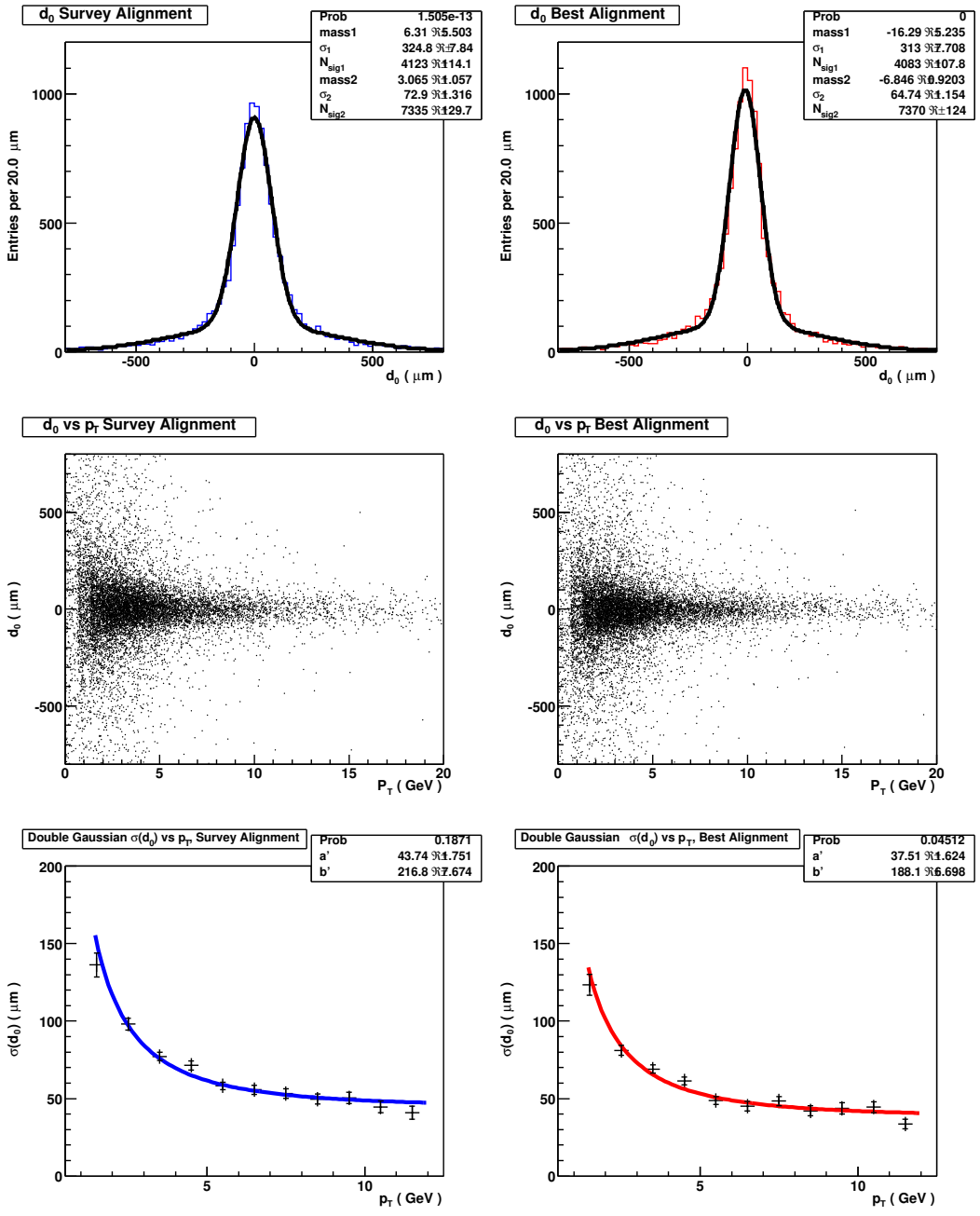

Figure 6.11: *The impact parameter distribution observed in CDF II data for tracks reconstructed with survey and best alignment constants. The width of the distribution $\sigma_{obs}(d_0)$ was estimated by sequential double Gaussian fits in slices of p_T. The fitted width of the narrowest Gaussian of each p_T slice as a function of p_T, were fitted by the function $\sigma_{obs}(d_0) = a' \oplus b'/p_T$ as shown.*

the survey and for the best alignment constants are plotted in Figure 6.11. The width of these *observed* impact parameter distributions $\sigma_{obs}(d_0)$ includes the contribution from the primary vertex distribution, $\sigma_{\rm BS}$, as discussed in Section 6.5.3. The track impact parameters were plotted as a function of the track p_T, as shown in the central two plots of Figure 6.11. $\sigma_{obs}(d_0)$ was estimated for sequential slices in p_T, by a double Gaussian fit. The fitted width of the narrowest Gaussian for each p_T slice is plotted as a function of p_T in the lower plots of Figure 6.11, for the survey and best alignments. A function of the form: $a' \oplus b'/p_T$ was fitted to these plots, yielding the fit parameters in Table 6.2.

6.5.3 Beam Spot Size

The primary vertex distribution, or beam spot size, $\sigma_{\rm BS}$, must be accounted for to find the impact parameter resolution, $\sigma(d_0)$, because these are combined to give the observed impact parameter distribution [Don01]:

$$\sigma_{obs}(d_0) = \sigma(d_0) \oplus \sigma_{\rm BS} \ [\mu m] \tag{6.13}$$

The $\sigma_{obs}(d_0)$ fit parameters, a' and b', can therefore be related to the $\sigma(d_0)$ parameters, a and b, in Equation 6.12 by:

$$a' \oplus b'/p_T = a \oplus b/p_T \oplus \sigma_{\rm BS} \tag{6.14}$$

Assuming a common multiple scattering term $b' = b$:

$$a = \sqrt{a'^2 - {\sigma_{\rm BS}}^2} = a' \ominus \sigma_{\rm BS} \tag{6.15}$$

Therefore the resolution term may be corrected if the beam spot size is known. Unfortunately, $\sigma_{\rm BS}$ is not yet well measured at CDF II, with estimates ranging from 25 μm to 35 μm; a default value of 30 μm is currently recommended [Pau04].

6.5.4 Misalignment Induced $\sigma(d_0)$ Degradation

Assuming $\sigma_{\rm BS} = 30$ μm, the corrections to the resolution term, a, are as shown in Table 6.2.

The ratio of the $\sigma_{\rm BS}$ corrected resolution terms for the survey and best alignments was found to be $F^{d_0}_{S/B} = a_S/a_B \approx 1.4$. The misalignment induced degradation factor is compatible with the factor, $F^{p_T}_{S/B}$, found for the p_T resolution in Section 6.3. A check of the sensitivity of $F^{d_0}_{S/B}$ to the $\sigma_{\rm BS}$ correction value is shown in Figure 6.12. The non-linearity in this plot creates an average degradation factor higher than that at $\sigma_{\rm BS} = 30$ μm. However, the degradation factor corresponding to this default beam spot size, $\sigma_{\rm BS} = 30$ μm, was taken in the following analysis.

Alignment	Fit parameters to width of impact parameter distribution vs p_T		Parameter a including beam spot correction	Misalignment Factor a_S/a_B
	a' $[\mu m]$	b' $[\mu m\, GeV]$	$a = a' \ominus \sigma_{\mathrm{BS}}$ $[\mu m]$	
Survey	43.74	216.8	31.8	1.4
Best	37.51	188.1	22.5	

Table 6.2: *The evaluated parameters in the fit of $a' \oplus b'/p_T$ to $\sigma_{obs}(d_0)$ versus p_T, of Figure 6.11. After correcting the fitted resolution terms $a'_{S,B}$ for a beam spot size of 30 μm, the misalignment induced degradation factor is $F^{d_0}_{S/B} = a_S/a_B \approx$ 1.4.*

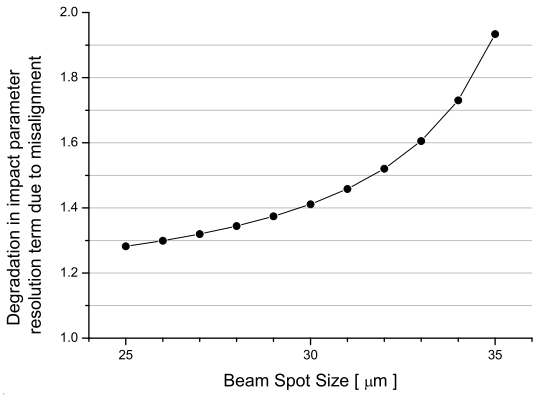

Figure 6.12: *The sensitivity of the misalignment induced degradation factor a_S/a_B in $\sigma(d_0)$, to variations in the beam spot size taken as a correction. The mean degradation factor of this plot is 1.48.*

Future work

Notably the fitted multiple scattering parameters have the ratio, $b'_S/b'_B = 1.15$, rather than being equal as expected. Why the multiple scattering term should apparently change with misalignment should be investigated in future work.

6.6 Effect of Impact Parameter Degradation

6.6.1 Introduction

The previous section concluded that a significant degradation factor, $F^{d_0}_{S/B} \approx$ 1.4, occurred in the impact parameter resolution of CDF II, due to misalignments during commissioning. In this section, the effect is mapped onto ATLAS and the implications for b-tagging are investigated.

Effect of misalignment on ATLAS impact parameter resolution

The ATLAS Inner Detector impact parameter resolution for the DC1 geometry has been parameterized as [Gar03a]:

$$\begin{aligned} \sigma^{\mathrm{LIN}}_{\mathrm{DC1}}(d_0) &= a + b/p_T & (6.16)\\ &= 11.7 + 106.8/p_T\ [\mu m] & (6.17)\end{aligned}$$

for d_0 in μm and p_T in GeV. This linear sum parameterization is not the same as the quadrature sum parameterization used in the above CDF II analysis and in the ATLAS TDR. However, at low and high p_T the parameterizations should be equivalent so to a good approximation the ATLAS DC1 parameterization is:

$$\begin{aligned} \sigma(d_0) &= a \oplus b/p_T & (6.18)\\ &\approx 11.7 \oplus 106.8/p_T\ [\mu m] & (6.19)\end{aligned}$$

When the misalignment factor, $F^{d_0}_{S/B} = 1.4(1)$, is applied, the resolution should be degraded during commissioning to:

$$\begin{aligned} \sigma_{COMS}(d_0) &= (F^{d_0}_{S/B} \times a) \oplus b/p_T & (6.20)\\ &\approx 16.5 \oplus 106.8/p_T\ [\mu m] & (6.21)\end{aligned}$$

In the following analysis, this misalignment effect was applied to ATLAS DC1 data that had already been reconstructed. The tracks were therefore already effectively smeared by $\sigma(d_0)$, so the impact parameter of each track had to be smeared by an additional term, a_{AL}, such that:

$$\begin{aligned} \sigma_{COMS}(d_0) &= a_{AL} \oplus \sigma(d_0)\\ &= a_{AL} \oplus a \oplus b/p_T & (6.22)\end{aligned}$$

Equating (6.20) with (6.22):

$$\begin{aligned} (F^{d_0}_{S/B} \times a) &= a_{AL} \oplus a\\ a_{AL} &= (F^{d_0}_{S/B} \times a) \ominus a\\ a_{AL} &\approx 11.6(5)\ [\mu m] & (6.23)\end{aligned}$$

The DC1 simulations included an estimated LHC beam spot size of $\sigma^{\mathrm{LHC}}_{\mathrm{BS}} = 15\ \mu$m, so the observed impact parameter distribution during commissioning should be:

$$\begin{aligned} \sigma_{obs}(d_0) &= a_{\mathrm{AL}} \oplus a \oplus b/p_T \oplus \sigma_{\mathrm{BS}}\\ &= 11.6(5) \oplus 11.7 \oplus 106.8/p_T \oplus 15\ [\mu m] & (6.24)\end{aligned}$$

6.6.2 *b*-tagging

The ability to identify *b*-jets is critical to many discovery channels at the LHC, in which new particles, such as Higgs bosons, couple preferentially to heavy quarks. The *b*-jets must be efficiently distinguished from the enormous light-jet and gluon-jet background. Various lifetime-based *b*-tagging algorithms have been developed with this aim and the algorithms continue to be enhanced. In this section, the degradation in the *b*-tagging performance due to misalignment is investigated for two benchmark algorithms.

b-tagging algorithms

The two *b*-tagging algorithms [ATL99, Gar03a] studied are based on the track impact parameter significance, $S = sign \times |d_0|/\sigma_{d_0}$, which assigns a higher weight to precisely measured tracks. The *sign* of the impact parameter significance is defined using the nearest jet direction, to be positive if the track crosses the jet-axis in front of the primary vertex, and negative otherwise.

On average, tracks from *b*-hadrons have large and positively signed impact parameters, while tracks from light-jets have impact parameters compatible with zero. The track significance distribution therefore exhibits a long tail at positive values for *b*-jet tracks, in contrast to the narrow distribution for light jets, as shown in Figure 6.13. This difference is exploited in the *b*-tagging method. The distributions both contain a central core described by a Gaus-

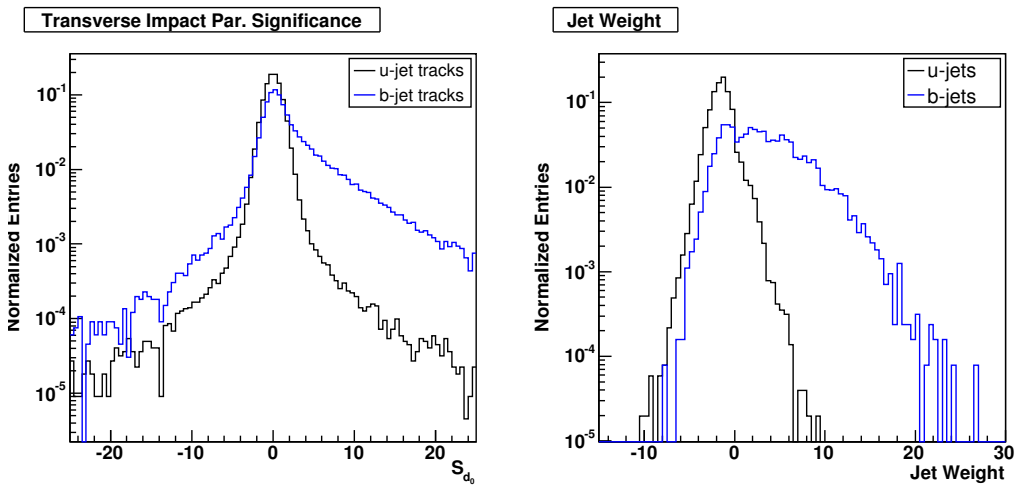

Figure 6.13: *The transverse impact parameter significance distributions for tracks in u-jets and b-jets and the corresponding jet weight for the 2D b-tagging algorithm.*

sian of width ∼1, representing correctly reconstructed tracks coming from the primary vertex. The light jets have a slight positive asymmetry arising from: long-lived particle decays (so called V^0's, e.g. Λ, K^0), daughters of heavy quarks formed in the fragmentation, and interactions with material.

The significance distributions are used to construct likelihood functions $f_b(S_t)$ and $f_u(S_t)$ for the track, t, to belong to a b-jet or u-jet. A weight is assigned to each track according to the ratio:

$$w_t = f_b(S_t)/f_u(S_t) \tag{6.25}$$

which represents the test statistic with which one may obtain the highest purity sample for a given signal efficiency, according to the *Neyman-Pearson lemma* [PDG02]. A jet weight is defined as the sum of logarithms of the track weights for all *good* tracks, i, in the jet:

$$w_{jet} = \sum_{i \in jet} \log w_t^i \tag{6.26}$$

The weight distributions for b- and u-jets are illustrated in Figure 6.13. Applying a cut on the jet weight yields a u-jet rejection factor $R_u = 1/\epsilon_u$, for a specified b-jet selection efficiency, ϵ_b.

The benchmark b-tagging algorithm of the ATLAS TDR, referred to here as the *2D algorithm*, uses only the transverse impact parameter, d_0, so that:

$$w_{jet}^{2D}(S_{d_0}) = \sum_{i \in jet} \log[f_b(S_{d_0})/f_u(S_{d_0})] \tag{6.27}$$

The b-tagging performance has been improved by combining the transverse and longitudinal impact parameter significances into a single weight:

$$w_{jet}^{3D}(S_{d_0}, S_{z_0}) = \sum_{i \in jet} \log[f_b(S_{d_0}, S_{z_0})/f_u(S_{d_0}, S_{z_0})] \tag{6.28}$$

The 2-dimensional function for this *3D algorithm* is shown in Figure 6.14. A binning in the (S_{d_0},S_{z_0}) plane, as defined elsewhere [Cor03], is currently used to compute the likelihood functions in the 3D method.

The impact of misalignment on b-tagging performance for the two algorithms was investigated.

6.6.3 Signal Generation and b-tagging Tools

Data sample

A large sample of $WH \rightarrow \mu\nu_\mu q\overline{q}$ events were generated, simulated and reconstructed for b-tagging studies in the ATLAS Data Challenge(DC1 phase 1)

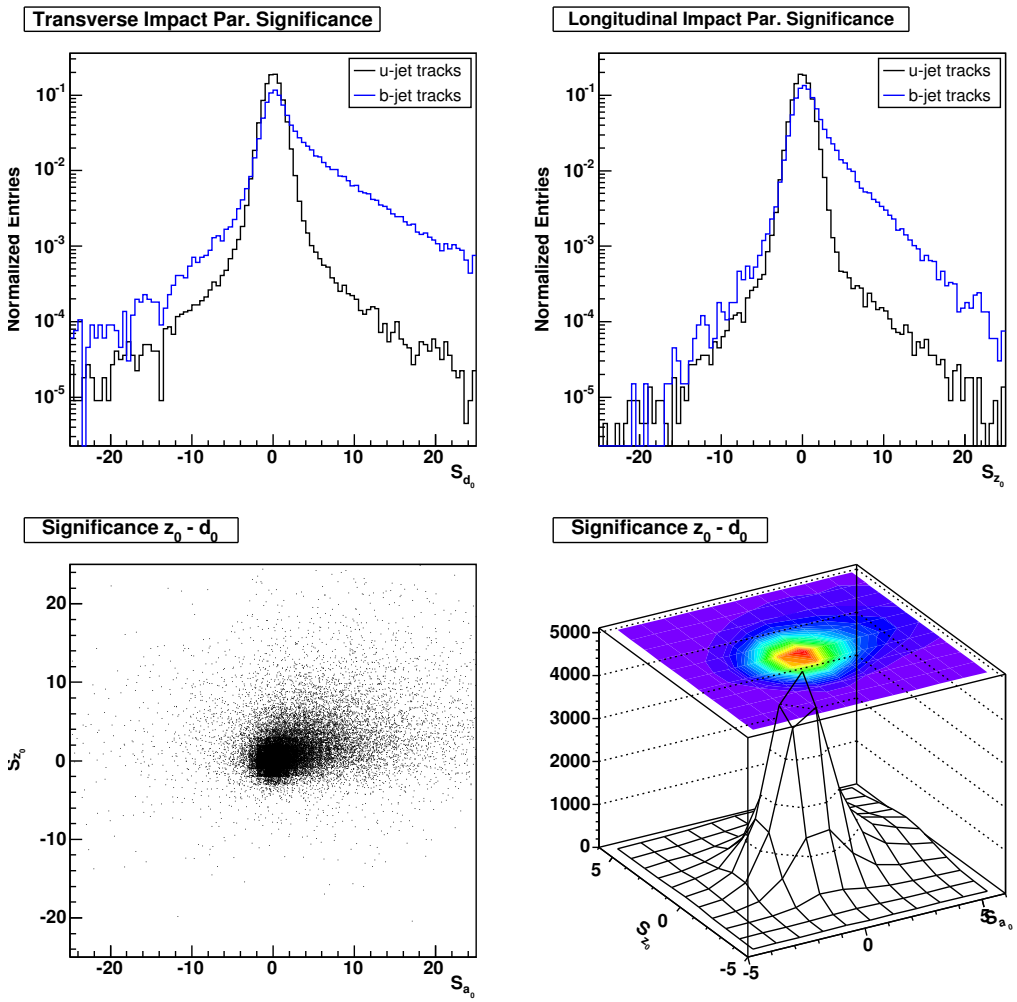

Figure 6.14: *The transverse and longitudinal impact parameter significances form a 2D dimensional track weight in the (S_{d_0},S_{z_0}) plane, from which the likelihood function is constructed for the 3D b-tagging method.*

framework [DC1]. In this study, 10k $b\bar{b}$ events and 40k $u\bar{u}$ events from DC1 were analysed, corresponding to a Standard Model Higgs boson with a mass of, $m_{\mathrm{H}} = 120$ GeV. The events were generated with PYTHIA 6.203. The detector response was simulated using DICE and GEANT 3.21, and the data were reconstructed with ATRECON, using xKalman++, as described elsewhere [Gar03a].

b-tagging code

The *b*-tagging algorithms outlined above were implemented in C++ code, that was developed from the original FORTRAN version written by E. Ros for the ATLAS TDR. The code was translated and updated by M. Vos for DC1 studies and most recently updated by J. E. Garcia [Gar03b].

6.6.4 Degradation of *b*-tagging Performance

The implications for *b*-tagging of commissioning misalignments, mapped to ATLAS from those observed at CDF II, are investigated in this section. The degradation in impact parameter resolution calculated in Section 6.6.1 was applied to the reconstructed tracks, by smearing the impact parameter of each track by the additional misalignment term, $a_{\mathrm{AL}} = 11.6(5)\ \mu$m, in Equation 6.24. The *b*-tagging performance was calculated with and without this additional smearing term, for the 2D and 3D *b*-tagging algorithms. The normalized track significance and weight for *u*- and *b*- jets in the 2D algorithm are plotted with and without the additional smearing term in Figure 6.15.

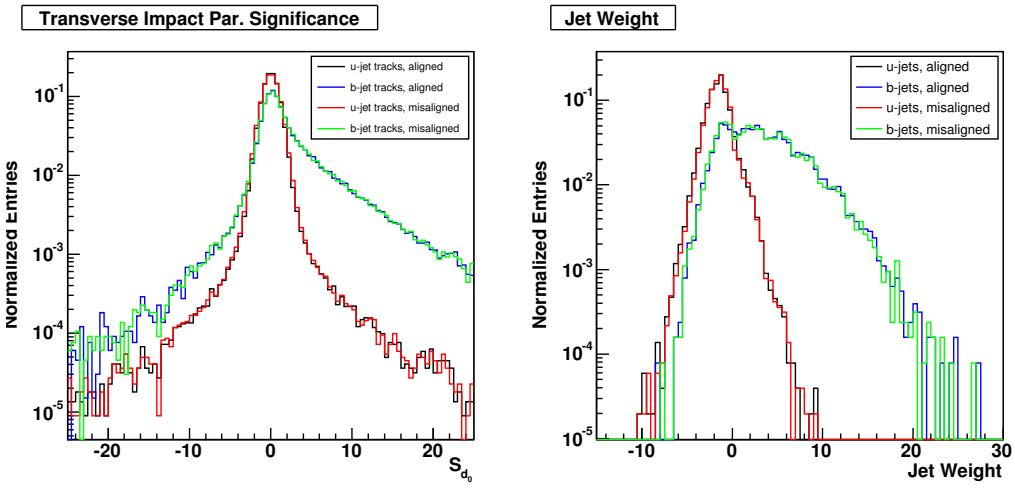

Figure 6.15: *Effect of misalignment on the transverse impact parameter significance distributions for tracks in u-jets and b-jets and the corresponding jet weight for the 2D b-tagging algorithm.*

The relation of the *u*-jet rejection, R_u, to the *b*-tag efficiency, ϵ_b, is plotted for both *b*-tagging algorithms in Figure 6.16. The *u*-jet rejections at the often quoted values of *b*-tag efficiency of 50% and 60% are presented in Table 6.3. The commissioning misalignment applied creates a degradation in *b*-tagging performance of around 10%.

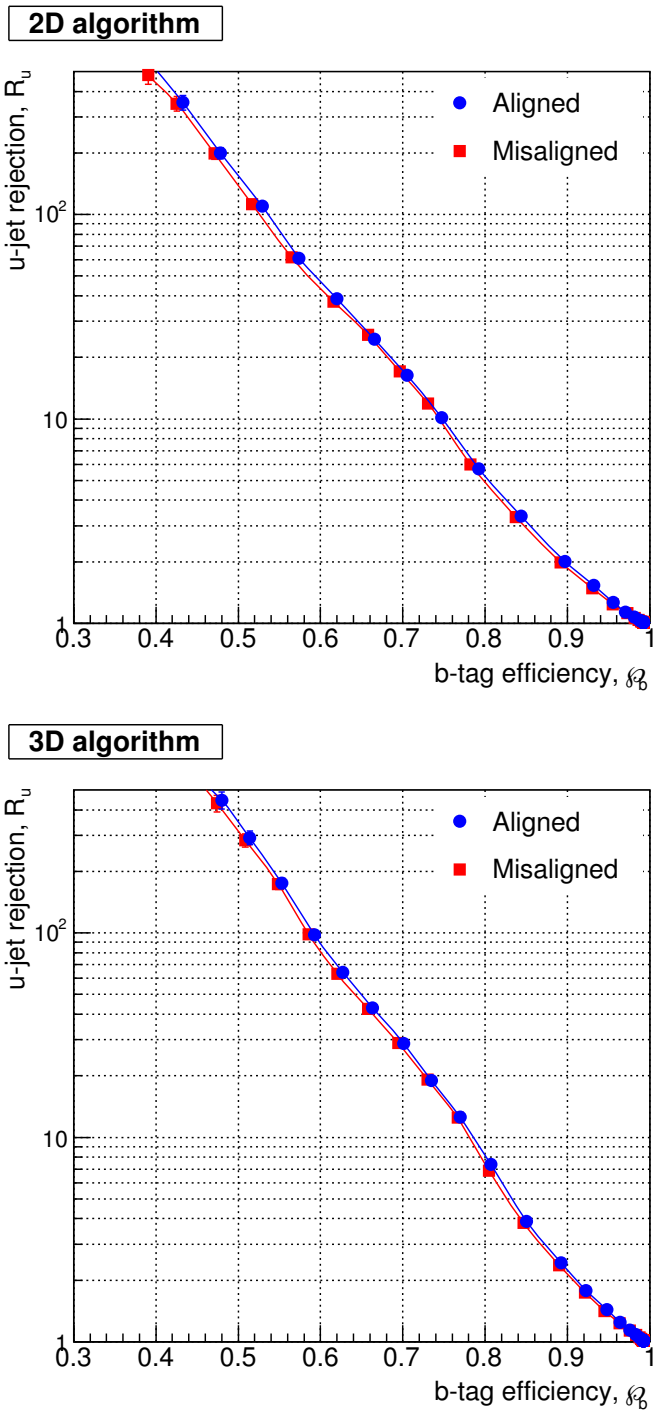

Figure 6.16: *b-tagging in ATLAS with and without misalignment. The u-jet rejection versus b-tag efficiency is shown for the 2D and 3D algorithms.*

		2D algorithm		*3D algorithm*	
		Perfect	Misaligned	Perfect	Misaligned
$R_u(\epsilon_b)$	$\epsilon_b = 50\%$	155 ± 8	138 ± 6	348 ± 22	316 ± 19
	$\epsilon_b = 60\%$	47.1 ± 1.1	43.6 ± 1.0	89.0 ± 3.4	81.7 ± 2.7
$\frac{R_u^{\text{Misaligned}}}{R_u^{\text{Perfect}}}$	$\epsilon_b = 50\%$	0.89		0.91	
	$\epsilon_b = 60\%$	0.92		0.92	

Table 6.3: *The effect of the predicted commissioning misalignment on b-tagging.*

		$\sigma_{R\phi} = 5$ μm $\sigma_Z = 15$ μm	$\sigma_{R\phi} = 10$ μm $\sigma_Z = 30$ μm	$\sigma_{R\phi} = 20$ μm $\sigma_Z = 60$ μm
2D $\frac{R_u^{\text{Misaligned}}}{R_u^{\text{Perfect}}}$	$\epsilon_b = 50\%$	0.95	0.89	0.65
	$\epsilon_b = 60\%$	0.98	0.91	0.74
3D $\frac{R_u^{\text{Misaligned}}}{R_u^{\text{Perfect}}}$	$\epsilon_b = 50\%$	0.99	0.91	0.67
	$\epsilon_b = 60\%$	0.97	0.92	0.71

Table 6.4: *The effect of random shifts of detector modules on b-tagging [Cor03].*

This result may be compared with a different b-tagging study [Cor03] in which random Gaussian shifts in Rϕ and Z of pixel and SCT detector modules were applied. This independent study used a slightly different algorithm to analyse $t\bar{t}H$ (120 GeV) and $t\bar{t}$ events, but the degradation *ratio* should be comparable to the study presented above, to first order. The degradation in u-jet rejections for different random misalignments found in that study are summarized in Table 6.4. On comparing the $\sim$10% degradation found for the above study, in Table 6.3 with those in Table 6.4, a match occurs for the case where the applied random misalignments were $\sigma_{R\phi} = 10$ μm and $\sigma_Z = 30$ μm.

The extent of the b-tagging degradation due to the predicted commissioning misalignment, is significant and shows the importance of reducing these misalignments as quickly as possible to increase the sensitivity of ATLAS to many discovery physics channels.

However, the degradation could be larger, because the misalignment observed at CDF II is thought to be due primarily to a sagitta deformation, as discussed in Section 6.3.4. In this case, the misalignment errors in a particular detector region should be correlated, giving rise to smaller errors in the measured impact parameter, due to the vertex constraint in the alignment procedure.

Also, the uncertainty of the beam spot size may have resulted in an underestimate of the degradation factor, given the sensitivity of the prediction to the beam spot size. As shown in Figure 6.12, the factor ranges from 1.3 to 1.9 for the beam spot sizes measured at CDF II. In future, the commissioning

misalignments predicted over this range could be applied systematically in these studies to evaluate the further degradation in b-tagging performance.

6.7 Conclusion

The effect of misalignments at CDF II on the transverse momentum and impact parameter resolutions have been studied. The resolution term in both of these track parameters was degraded by $\sim$1.4 for the initial alignment compared to the current best alignment. These effects were mapped onto the ATLAS detector to predict the track parameter degradation during detector commissioning.

The implications for ATLAS physics have been examined. The effect of momentum degradation on low $p_T(\lesssim 30$ GeV) physics should be negligible at ATLAS, due to the dominance of multiple scattering. Higher p_T channels are expected to be significantly degraded. The predicted degradation in the impact parameter resolution caused a 10% degradation in the b-tagging performance.

A substantial source of the misalignment at CDF II is thought to be due to a sagitta misalignment. The importance of quickly reducing the misalignments, to increase the sensitivity of ATLAS to discovery and precision physics has been emphasized by these studies. The SCT alignment system should be a valuable tool in ATLAS to quicken the commissioning phase and improve the ultimate alignment precision.

Chapter 7

Conclusion

A novel alignment system has been developed for the ATLAS SemiConductor Tracker. This system will directly measure short time-scale motions and complex deformations of the ATLAS SCT, which are inaccessible with and also degrade, track based alignment methods. This system forms part of a coherent alignment strategy that will rapidly enhance the sensitivity of ATLAS to discovery and precision physics. This thesis contributes to the development of this alignment system.

The alignment system is based on a geodetic grid of length measurements between nodes attached to the SCT. These length measurements within ATLAS are possible by the specially developed technique of Frequency Scanning Interferometry. This work contributes important advances in the FSI technique, which improve the reliability, precision and scope of the measurements. These improvements enabled a 170 mm stabilized reference interferometer to be measured with a precision of 7 nm (0.04 ppm).

The FSI measurements were shown to be tolerant to several types of spurious effects that might degrade the interferometer signals in ATLAS. Methods to suppress unwanted spurious reflections were developed and incorporated in the design of ATLAS FSI components.

The first, prototype FSI geodetic grids have been constructed as part of a prototype ATLAS FSI system, that has demonstrated remote, multiple, simultaneous length measurements between grid nodes. Precise length measurements were possible with FSI, despite the low power signals ($\sim$10 pW) inherent with the fibre-coupled, zero-maintenance interferometer design. Methods to optimize the interferometer signals were developed. Some optical cross-talk was observed between interferometers, but this had negligible effect on the χ^2 profile, which indicates the quality of a sine fit to the interferometer signals. The lengths of robust, fibre-couple interferometers, which had compact components of the same design as for ATLAS, were measured with an average precision of

$\sim$200 nm; well within the required precision of 1 μm.

The ATLAS FSI system will monitor shape deformations of the SCT, as reconstructed from FSI measurements. Investigations of shape reconstruction were performed with the prototype grids. A reconstruction method was developed, that was checked by Monte Carlo studies and with data from the prototype grids. Fine adjustments to the prototype grid shape were reconstructed very precisely: by a factor of $\gtrsim$4 better than the required precision. Independent measurements verified the node reconstruction was linear to within the resolution of the measurements. Complex patterns of node movements, were reconstructed in three dimensions from single FSI grid measurements at each node position, as required for the ATLAS FSI system.

The reconstruction process was checked by exploiting the redundancy in the prototype grids. The systematic removal of a redundant measurement from the grid in the analysis had almost no effect on the mean reconstructed position of the grid nodes, after an appropriate calibration of the model. A similar calibration will be crucial in ATLAS to prevent grid reconstruction errors if one or more redundant grid lines fail. The reconstruction models should also include estimated length measurement precisions, for the best precision on the node co-ordinates. The simulated rigidity of the prototype grids closely matched that found from the repeated measurements of the prototype grids, indicating the simulation and reconstruction methods are sound.

The FSI grids of the ATLAS SCT were simulated in detail. In general the simulations indicate the FSI system should be capable of reconstructing displacements of the grid nodes to well within the required precision. Certain assumptions on the rigidity of the support structures must be met for the predicted precisions to be valid. FEA models indicate this should be the case.

An FSI grid for the SCT End-cap was developed in this work and optimized for rigidity and redundancy. The future challenge will be to combine data from the FSI system and from tracks, and incorporate knowledge of the support structures, to align the ATLAS Inner Detector. A programme to guide these future developments has been suggested.

The magnitude of commissioning misalignments at CDF II were studied and mapped onto the ATLAS detector. The intrinsic resolution term in the both the transverse momentum and impact parameter resolutions was degraded by $\sim$1.4. The implications are that for low p_T physics in ATLAS, the effect of misalignment induced momentum degradation should be negligible, but is expected to be significant for $p_T > 30$ GeV. The b-tagging performance of the commissioning alignment was degraded by 10%. The vital role of the SCT alignment system is emphasized by these studies.

References

[ALI99] ALICE Collaboration, *ALICE Technical Design Report of the Inner Tracking System.* CERN/LHCC99-12, June 1999.

[ATL97] ATLAS Collaboration, *ATLAS Inner Detector Technical Design Report.* CERN-LHCC/97-16, April 1997.

[ATL98] ATLAS Collaboration, *Trigger Performance Status Report.* CERN-LHCC/98-015, August 1998.

[ATL99] ATLAS Collaboration, *ATLAS Detector and Physics Performance Technical Design Report.* CERN-LHCC/99-(14/15), May 1999.

[Bai01] J. Baines et al, *Effects of Inner Detector Misalignment and Inefficiency on the ATLAS B-physics Trigger.* ATL-DAQ-2001-006.

[Bar99] D. Barberis, *ATL-CONF-99-03: ATLAS Inner Detector developments.* CERN, 1999.

[Bar01] G. Barbier, *Engineering drawings. ATLAS SCT Barrels: B3 252650P0(1-4), B4 252651P0(1-4), B5 252652P0(1-4), B6 252653P0(1-4). Mod. 2.* University of Geneva, 2001.

[Bre91] A. Breakstone, *Performance of the capacitive displacement measuring system of the Mark II detector at the SLC.* NIM A **281** 453-461 (1989).

[Bro02] N. Brooks et al., *Engineering drawings. FSI Scorpion Design N214-008. FSI Endjewel Design N214-009.* Physics Design Office, University of Oxford, UK, 2002.

[Bru98] L. Brunel, *SIMULGEO Simulation and reconstruction software for opto-geometrical systems.* CERN, CMS Note 1998/079. `http://cmsdoc.cern.ch/cgi-doc/Grep_Search_F/documents/allxx.html?STRING=NOTE1998079`

[Bur01] K. Burkett on behalf of the CDF Collaboration, *Design and construction of the CDF central outer tracker.* NIM A **461** 62-64 (2001).

[Cac92] M.Caccia et al. *1990 data taking period results of the DELPHI CDM System.* NIM A **315** 143-148 (1992).

[CDF94] CDF Collaboration, *Evidence for top quark production in $p\overline{p}$ collisions at $\sqrt{s} = 1.8$ TeV.* Phys. Rev. D **50**(5) 2966 (1994).

[CDF96a] CDF Collaboration, *Measurement of the Mass of the B_s^0 Meson.* Phys. Rev. D **53**(7) 3496 (1996).

[CDF96b] CDF Collaboration, *The CDF II Detector Technical Design Report.* FERMILAB-Pub-96/390-E.

[Coe01] P. A. Coe, *An Investigation of Frequency Scanning Interferometry for the alignment of the ATLAS semiconductor tracker.* Doctoral Thesis, University of Oxford, UK, 2001. `http://www-pnp.physics.ox.ac.uk/~atlas/Documents.html#Theses`

[Coe02a] P. A. Coe, *ATLAS SCT Alignment On-detector components. PRR* ATL-IS-ES-0074, University of Oxford, UK, 2002. `https://edms.cern.ch/document/348798/1`

[Coe02b] P. A. Coe, S. M. Gibson, *ATLAS SCT Alignment Review of Frequency Scanning Interferometry. PRR* ATL-IS-ES-0075, University of Oxford, UK, 2002. `https://edms.cern.ch/document/348799/1`

[Coe04] P. A. Coe, *Private Communication.* University of Oxford, UK, 2004.

[Cor03] S. Corréard et al, *b-tagging with DC1 data.* ATL-COM-PHYS-2003-049. November 2003.

[Cug01] J. Cugnoni, *ATLAS SCT Assembly Finite Element Analysis.* University of Geneva. 18 June 2001.

[Dau94a] C. Daum, *Alignment of the ATLAS Inner Detector.* ATLAS Collaboration Internal Note INDET-NO-060, 12 September 1994.

[Dau94b] C. Daum, *Alignment of the ATLAS detector from Inner Detector to Muon Spectrometer.* ATLAS Collaboration Internal Note TECH-NO-013, 17 February 1994.

[Dau95] C. Daum, *Least Square Fit and Covariance Matrix for and Octagonal Alignment Grid for the ATLAS Detector.* ATLAS Collaboration Internal Note TECH-NO-014 Revised Version, 23 February 1995.

[DC1] ATLAS Data Challenge 1, phase 1. Report to be published, draft available at: `http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/DC/DC1/`

[Don01] S. Donati and L. Ristori, *Measuring Beam Width and SVX Impact Parameter Resolution.* CDF/MEMO/TRACKING/CDFR/4189. (Revised 2001).

[Dou01] G. Doucas, J. Grosse-Knetter, R. Nickerson and L. Vertogradov *X-ray tomography for the ATLAS semi-conductor tracker.* NIM A **457** 43-51 (2001).

[Due01] I. Duerdoth et al, *Measurements of convection currents between the disks of the ATLAS SCT end cap.* ATL-IS-TR-0010. Version 1. (5 October 2001).

[Ell98] J. R. Ellis, *Beyond the Standard Model for Hillwalkers.* hep-ph/9812235.

[Fibre] Fibercore Limited QA certificate YD502-03. SM800 optical fibre parameters:
Fibre OD: 125 μm
Numerical aperture: 0.1
Cut-off wavelength: 756 nm
Attenuation: 2 dB/km at 830 nm
Optimum Launch spot-size: 7 μm

[Fos99] I. Foster and C. Kesselman, editors, *The Grid: Blueprint for a new computing infrastructure.* Morgan Kauffman Publishing Inc., 1999.

[Fox96] A. F. Fox-Murphy, *Development of a novel alignment system for the ATLAS inner detector and an investigation of the effect of alignment inaccuracies on tracker performance.* Doctoral Thesis, University of Oxford, UK, 1996. `http://www-pnp.physics.ox.ac.uk/~atlas/Documents.html#Theses`

[Gad00] S. Gadomski, *Updated impact parameter resolutions of the ATLAS Inner Detector.* ATL-INDET-2000-020.

[Gar03a] J. E. Garcia Navarro et al, *The b-tagging performance of the complete ATLAS DC1 layout using WH events.* ATL-COM-INDET-2003-017.

[Gar03b] J. E. Garcia Navarro, *Private Communication.* 2003.

[Gib01a] S. M. Gibson, *ATLAS SCT Alignment, Light path protection tubes.* ATL-IS-ES-0031, University of Oxford, UK, 2001. `https://edms.cern.ch/document/312457/1`

[Gib01b] S. M. Gibson, *ATLAS SCT Alignment, Forward Grid Simulations.* ATL-IS-AP-0054, University of Oxford, UK, 2001. https://edms.cern.ch/document/324552/1

[Gib03] S. M. Gibson and A. J. M. Muijs, *ATLAS SCT End-cap Alignment System Layout.* ATL-IS-ES-0081, University of Oxford, UK, 2003. https://edms.cern.ch/document/377555/2

[Gol02] D. Goldstein and D. Saltzberg, *The RASNIK Real-Time Relative Alignment Monitor for the CDF Inner Tracking Detectors.* NIM A **506**(1-2) 92-100 (2003).

[Gre99] A. S. Grewal, *Detector Development for ATLAS and Supersymmetry Physics Studies.* Doctoral Thesis, University of Oxford, UK, 1999. http://www-pnp.physics.ox.ac.uk/~atlas/Documents.html#Theses

[Har91] K. Harvey & C. Myatt, *External-cavity diode laser using a grazing-incidence diffraction grating.* Optics Letters, **16** 910-912 (1991).

[Hay00] S. Haywood et al., *Offline Alignment and Calibration of the Inner Detector.* ATLAS-INDET-2000-005.

[Hay02] S. Haywood et al., *Determination of SCT Wheel Positions.* ATLAS-IS-ES-0080. Rutherford Appleton Laboratory, 2002. https://edms.cern.ch/document/361723/1

[Haw04] R. J. Hawkings, on behalf of the ATLAS collaboration, *The ATLAS inner detector and flavour tagging performance.* DOI 10.1007/s1010504a123, EPJdirect A**1** 1-11 (2004).

[Hec87] E. Hecht, *Optics (2nd Ed.).* Addison-Wesley, Reading Massachusetts, 1987.

[Hin04] D. Hindson, *A Robust Procedure for the Alignment of the ATLAS Inner Detector Using Tracks.* ATL-COM-INDET-2004-005.

[Hol66] James G. Holbrook, *Laplace Transforms for Electronic Engineers.* Oxford: Pergamon Press, 1966.

[How01] D. F. Howell et al., *ATLAS-SCT-Alignment Overview.* ATL-IS-ES-0026, University of Oxford, UK, 2001.

[Huf99] Original ATLAS SCT End-cap alignment grid design. http://www-pnp.physics.ox.ac.uk/~atlas/alignment/SCT/forward_main.html

[Huf01] WHIPM: WHite light Interferometric Profile Measurement system. Developed at the University of Oxford. B. Todd Huffman, *ATLAS SCT - Alignment - Retro-reflectors.* ATL-IS-ES-0029, University of Oxford, UK, 2001.

[Huf02] B. Todd Huffman, *Private Communication.* University of Oxford, UK, February 2002.

[Jam75] F. James and M. Roos, *Minuit - a system for function minimization and analysis of the parameter errors and correlations.* Computer Physics Communications, **10**(6) 343-367 (December 1975).

[Kay89] G. W. C. Kaye and T. H. Laby, *Tables of Physical and Chemical Constants.* Longman Scientific & Technical, Harlow England, 1986 (Reprinted 1989).

[Kua96] S. Kuang, *Geodetic Network Analysis and Optimal Design: Concepts and Applications.* Chelsea, Mich. : Ann Arbor Press, 1996.

[Lau01] W. Lau, *Private Communication.* University of Oxford, UK, 2001.

[LEPEWG] Preliminary results from LEP Electroweak Working Group, Summer 2003. `http://www.cern.ch/LEPEWWG/`

[LHCb03] LHCb Collaboration, *LHCb Reoptimised Detector Design and Performance TDR.* CERN-LHCC-2003-030, September 2003.

[Mar77] D. Marcuse, *Loss Analysis of Single-Mode Fiber Splices.* The Bell System Technical Journal, **56**(5) 703-718 (May-June 1977).

[Mat01] T. Matsushita et al, *Optical alignment system for the ZEUS micro vertex detector.* NIM A **466**(2) 383-389 (2001).

[McN01] R. McNulty et al, *A Procedure for the Software Alignment of the CDF Silicon System.* CDF/DOC/TRACKING/GROUP/5700. 8 August 2001.

[Mit02] A. Mitra, *Development of an APD amplifier for the runtime alignment of the ATLAS Inner Detector and SUSY studies.* Doctoral Thesis, University of Oxford, UK, 2002. `http://www-pnp.physics.ox.ac.uk/~atlas/Documents.html#Theses`

[Mit03] A. Mitra, *Private Communication.* University of Oxford, UK, 2003.

[Mits03] V. A. Mitsou, on behalf of the ATLAS TRT collaboration, *The ATLAS Transition Radiation Tracker.* ATL-IT-AM-0041 and proceedings from 8th ICATPP, Como, Italy, 6-10 October 2003.

[New03] Newport breadboard surface flatness specification from: `http://www.newport.com`

[Nic96] R. B. Nickerson et al., *Frequency Scanned Interferometry (FSI): The basis of a survey system for ATLAS using fast automated remote interferometry.* NIM A **383** 229-237 (1996).

[Ost01] A. Ostapchouk et al, *The Alignment of the CMS Tracker.* CMS Note 2001/053.

[Pau03] Thilo Pauly, *Private Communication.* University of Oxford, UK, 2003.

[Pau04] Thilo Pauly, *Measurement of the Lifetime Difference of the B_S Meson at CDF using $B_S \rightarrow J/\psi\ \phi$.* Doctoral Thesis, University of Oxford, UK, 2004.

[PDG02] Particle Data Group, K. Hagiwara et al. *Review of Particle Physics.* Phys. Rev. D **66** 010001 (2002). `http://pdg.lbl.gov`

[Quig02] C. Quigg, *The Electroweak Theory.* arXiv:hep-ph/0204104v1, 2002.

[Qui02] Quill beam-splitter design, *Engineering Drawing. ATLAS quill light beam fused silica splitter, N214-013 Revision D.* Physics Design Office, University of Oxford, August, 2002.

[Reic01] A. Reichold, *ATLAS SCT - Alignment - Quill Test Specification.* ATL-IS-ES-0028, University of Oxford 2001.

[Ric98] E. Richter-Was, D. Froidevaux and L. Poggioli, *ATLFAST 2.0 a fast simulation package for ATLAS.* ATL-PHYS-098-131. 1998.

[Sal98] D. Saltzberg and E. Sanders, *Proposal for a Rasnik Relative Alignment Monitor for the CDF II SVX-ISL-COT.* CDF Internal Note 4510, March 1998.

[Sgu99] G. Sguazzoni et al, *Monitoring the stability of the ALEPH vertex detector.* Nucl. Phys. B (Proc. Sup.) **78** 301-306 (1999).

[Sil00] A. Sill on behalf of the CDF Collaboration, *CDF Run II Silicon Tracking Projects.* NIM-A **447** 1-8 (2000).

[Sjo01] T. Sjostrand et al, *High-energy-physics event generation with PYTHIA 6.1.* Computer Physics Commun. **135** 238 (2001).

[SM] S.L. Glashow, *Nucl. Phys.* **22** 579 (1961).

S. Weinberg, *Phys. Rev. Lett.* **19** 1264 (1967).

A. Salam, *Proc. 8th Nobel Symposium, Stockholm 1968, ed. N. Svartholm (Almqvist and Wiksells, Stockholm, 1968), p367.*

P. W. Higgs, *Phys. Rev. Lett.* **13** 508 (1964).

I. J. R. Aitchison and A. J. G. Hey, *Gauge Theories in Particle Physics. (third edition, Volume 1)* IOP Publishing, UK, 2002.

D. Griffiths, *Introduction to Elementary Particles.* Wiley, 1987.

F. Halzen and A. D. Martin, *Quarks and Leptons.* Wiley, 1984.

D. H. Perkins 2000, *Introduction to High Energy Physics. (fourth edition)* Cambridge University Press, 2000.

[Snip03] Q. H. C. Snippe, *Structural Analysis of the SCT End-cap Discs.* ATL-IS-ES-0004-D, NIKHEF, NL, 2003. `https://edms.cern.ch/document/325743/2`

[SNO] Q. R. Ahmad et al, *Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory. Phys. Rev. Lett.* SNO NC PRL, submitted for publication, 2001.

Q. R. Ahmad et al, *Phys. Rev. Lett.* **87** 071301 (2001).

[SuperK] S. Fukuda et. al. *Phys. Rev. Lett.* **86** 5651-5655 (2001).

[Tur01] M. Turala et. al. *The ATLAS Semiconductor Tracker* NIM-A **466** 243-254 (2001).

[Wei97] A. Weidberg and S. Snow, *The effect of inner detector misalignment on track resolution.* ATL-INDET-97-160, CERN, 1997.

[Wer03] P. Werneke et al., *ATLAS SCT End-cap disc fixing.* ATL-IS-ER-0023-C, NIKHEF, NL, 2003. `https://edms.cern.ch/document/324144/2`

[Wil03] I. Wilmut, *Engineering drawings of the SCT End-cap discs and alignment system. 0-TB-0049-177-01-M and TD-1013-200-asm and various concept drawings.* Rutherford Appleton Laboratory, UK, 2003.